

**TMA947/MMG621**  
**NONLINEAR OPTIMISATION**

**Date:** 21-01-02  
**Time:** 8<sup>30</sup>-13<sup>30</sup>  
**Aids:** All aids are allowed, but cooperation is not allowed  
**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:** Ann-Brith Strömberg

**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.*

### Question 1

(Simplex method)

Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) := -4x_1 + x_2, \\ \text{subject to} \quad & x_1 - x_2 \leq 2, \\ & -x_1 + 2x_2 \leq 1, \\ & x_1, \quad x_2 \geq 0. \end{aligned}$$

- (0.5p) a) Formulate the problem on the standard form for linear optimization problems.
- (1.5p) b) Solve the problem using the simplex method. Present an optimal solution in the original variables.
- (1p) c) Consider modifying the problem by including the variable  $x_3$  as follows

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) := -4x_1 + x_2 + x_3, \\ \text{subject to} \quad & x_1 - x_2 + x_3 \leq 2, \\ & -x_1 + 2x_2 - 3x_3 \leq 1, \\ & x_1, \quad x_2, \quad x_3 \geq 0. \end{aligned}$$

Solve the problem using the simplex method using the optimal basis from b) as initial basis. Present an optimal solution or a ray of unboundedness in the original variables

---

### (3p) Question 2

(Farkas Lemma)

Let  $B, C \in \mathbb{R}^{m \times n}$  be matrices and  $\mathbf{v} \in \mathbb{R}^m$  a vector. Assume that there exists a vector  $\mathbf{z} \leq \mathbf{0}^n$  such that

$$B\mathbf{z} = C\mathbf{z} + \mathbf{v}.$$

Show that there cannot exist a vector  $\mathbf{y} \in \mathbb{R}^m$  such  $\mathbf{v}^T \mathbf{y} > 0$  and  $C^T \mathbf{y} \leq B^T \mathbf{y}$ .

---

### Question 3

(KKT conditions)

Consider the following optimization problem, where  $\mathbf{c}$  is a nonzero vector in  $\mathbb{R}^n$ :

$$\begin{aligned} \max \quad & \mathbf{c}^T \mathbf{x}, \\ \text{s.t.} \quad & \mathbf{x}^T \mathbf{x} \leq 1. \end{aligned}$$

- (1p) a) Show that  $\bar{\mathbf{x}} = \mathbf{c}/\|\mathbf{c}\|$  is a KKT point.  
(2p) b) Show that  $\bar{\mathbf{x}}$  is the unique global optimal solution.
- 

### (3p) Question 4

(Gradient projection)

The gradient projection algorithm is a generalization of the steepest descent method to constrained optimization problems over convex sets. Given a feasible point  $\mathbf{x}^k$ , the next point is obtained according to  $\mathbf{x}^{k+1} = \text{Proj}_X(\mathbf{x}^k - \alpha_k \nabla f(\mathbf{x}^k))$ , where  $X$  is the convex set over which we minimize,  $\alpha_k > 0$  is the step length, and  $\text{Proj}_X(\mathbf{y}) = \arg \min_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|$ .

Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) = x_1^2 + 2x_2^2 - 2x_1x_2 - 2x_1 - 3x_2 + 8, \\ \text{subject to} \quad & \mathbf{x} \in X, \end{aligned}$$

where  $X$  is the rectangle  $X = \{\mathbf{x} \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 3 \text{ and } 0 \leq x_2 \leq 2\}$

Start at the point  $\mathbf{x}^0 = (0, 0)^T$  and perform two iterations of the gradient projection algorithm using step lengths  $\alpha_k = 1$  for all  $k$ . You may solve the projection problem in the algorithm graphically. Is the point obtained a global/local minimum? Motivate why/why not.

---

**(3p) Question 5**

(Modelling)

Consider a Sudoku, i.e., a  $3 \times 3$  matrix of cells where each cell is a  $3 \times 3$  matrix of tiles; the Sudoku thus forms a  $9 \times 9$  matrix of tiles. Each tile is to be assigned a number from one to nine such that the number is unique in the row, column, and cell containing the tile. The numbers of some tiles are given; an example of a Sudoku is illustrated in Figure 1.

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Figure 1: A Sudoku.

- (1.5p)** a) Create a binary linear model of the feasible assignments of a Sudoku. Let  $x_{ijk}$  denote the binary decision choice of assigning number  $k$  to row  $i$  and column  $j$ , where  $i, j, k \in \{1, \dots, 9\}$ . Let  $(i, j, k) \in A$  denote the set of initially given numbers, i.e.,  $x_{ijk} = 1$  for all  $(i, j, k) \in A$ .

*hint:* Introduce the sets  $C_l$  containing the tiles  $(i, j)$  in cell  $l$ ,  $l = 1, \dots, 9$ .

- (1.5p)** b) Assume that the Sudoku has a feasible solution  $\bar{\mathbf{x}}$ . Add a linear objective function to your model in a) such that  $\bar{\mathbf{x}}$  is an optimal solution if and only if it is the only feasible solution. Show that any other feasible solution  $\tilde{\mathbf{x}} \neq \bar{\mathbf{x}}$  has a better objective value.

**Question 6**

(true or false)

Indicate for each of the following three statements whether it is true or false. Motivate your answers!

- (1p)** a) Let  $S$  be a nonempty, closed and convex set in  $\mathbb{R}^n$ , and let  $f : \mathbb{R}^n \mapsto \mathbb{R}$  be defined as  $f(\mathbf{y}) = \min_{\mathbf{x} \in S} \|\mathbf{y} - \mathbf{x}\|$ .  
*Claim: The function  $f$  is convex.*
- (1p)** b) *Claim: If the KKT conditions are sufficient, then they are also necessary.*
- (1p)** c) *Claim: For the phase I (when a BFS is not known a priori) problem of the simplex algorithm, the optimal value is always zero.*

**(3p) Question 7**

(Lagrangian relaxation and decomposition)

Consider the problem to

$$\text{minimize} \quad z, \tag{1}$$

$$\text{subject to} \quad \sum_{j \in \mathcal{J}} p_{ij} x_{ij} \leq z, \quad i \in \mathcal{I}, \tag{2}$$

$$\sum_{i \in \mathcal{I}} x_{ij} = 1, \quad i \in \mathcal{J}, \tag{3}$$

$$x_{ij} \in \{0, 1\}, \quad j \in \mathcal{I}, j \in \mathcal{J}, \tag{4}$$

$$z \in \mathbb{R}. \tag{5}$$

Here  $\mathcal{I}$  denotes a set of machines and  $\mathcal{J}$  denotes a set of tasks,  $x_{ij}$  denotes the decision to perform task  $j$  by machine  $i$ , and  $p_{ij}$  denotes the corresponding processing time. The variable  $z$  denotes the makespan, i.e., the time at which the last machine is finished.

- (1p)** a) Lagrangian relax constraints (2) with multipliers  $u_i, i \in \mathcal{I}$ . Let  $h(\mathbf{u})$  denote the value of the dual function and show that  $h(\mathbf{u}) = -\infty$  if  $\sum_{i \in \mathcal{I}} u_i \neq 1$ .
- (1.5p)** b) Assume that  $\sum_{i \in \mathcal{I}} \bar{u}_i = 1$  and show that evaluating  $h(\bar{\mathbf{u}})$  reduces to solving  $\mathcal{J}$  separate optimization problems. State the optimal solution to each of these Lagrangian subproblems and the resulting formula for  $h(\bar{\mathbf{u}})$ .
- (0.5p)** c) Show that the Lagrangian subproblem solution forms a primal feasible solution for some value of  $z$ .
-

**TMA947/MMG621  
NONLINEAR OPTIMISATION**

**Date:** 21-01-02

**Examiner:** Ann-Brith Strömberg

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

## Question 1

(Simplex method)

(0.5p) a) The problem on standard form is:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) := -4x_1 + x_2, \\ \text{subject to} \quad & x_1 - x_2 + s_1 = 2, \\ & -x_1 + 2x_2 + s_2 = 1, \\ & x_1, x_2, s_1, s_2 \geq 0. \end{aligned}$$

(1.5p) b) We can start directly in phase two since the slack variables provides an initial feasible basis.

First iteration: we have  $x_B = (s_1, s_2), x_N = (x_1, x_2), B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,

$$N = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}, c_B = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, c_N^T = [-4 \quad 1], B^{-1}b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Checking optimality:

$$\bar{c}_N^T = c_N^T - c_B^T B^{-1}N = [-4 \quad 1] - [0 \quad 0] \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = [-4 \quad 1]$$

Not optimal, minimum reduce costs indicate  $x_1$  enter the basis.

Minimum ratio test:  $B^{-1}N_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$\operatorname{argmin}_{i \in (B^{-1}N_1)_i > 0} \frac{(B^{-1}b)_i}{(B^{-1}N_1)_i} = \operatorname{argmin}\left\{\frac{2}{1}, -\right\}$$

hence,  $s_1$  leaves the basis.

Second iteration: we have  $x_B = (x_1, s_2), x_N = (x_2, s_1), B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, N = \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}, c_B = \begin{bmatrix} -4 \\ 0 \end{bmatrix}, c_N^T = [1 \quad 0], B^{-1}b = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$

Checking optimality:

$$\bar{c}_N^T = c_N^T - c_B^T B^{-1}N = [1 \quad 0] + [4 \quad 0] \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix} = [-3 \quad 4]$$

Not optimal, minimum reduce costs indicate  $x_2$  enter the basis.

Minimum ratio test:  $B^{-1}N_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$\operatorname{argmin}_{i \in (B^{-1}N_2)_i > 0} \frac{(B^{-1}b)_i}{(B^{-1}N_2)_i} = \operatorname{argmin}\left\{-, \frac{3}{1}\right\}$$

hence,  $s_2$  leaves the basis.

Third iteration: we have  $x_B = (x_1, x_2)$ ,  $x_N = (s_1, s_2)$ ,  $B = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ ,  $B^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ ,  $N = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $c_B = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$ ,  $c_N^T = [0 \ 0]$ ,  $B^{-1}b = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ .

Checking optimality:

$$\bar{c}_N^T = c_N^T - c_B^T B^{-1} N = [0 \ 0] + [7 \ 3] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [7 \ 3] \geq 0$$

The solution in the original variables are  $x_1 = 5$ ,  $x_2 = 3$ .

(1p) c) Continuing the third iteration, we have a new non-basic variable  $x_3$ .

$$x_N = (x_3, s_1, s_2), N = \begin{bmatrix} 1 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}, c_N^T = [1 \ 0 \ 0].$$

Checking optimality:

$$\bar{c}_N^T = c_N^T - c_B^T B^{-1} N = [1 \ 0 \ 0] + [7 \ 3] \begin{bmatrix} 1 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} = [-1 \ 7 \ 3]$$

Not optimal, minimum reduce costs indicate  $x_3$  enter the basis.

Minimum ratio test:  $B^{-1}N_1 = \begin{bmatrix} -1 \\ -2 \end{bmatrix} \leq 0$ , hence the problem is unbounded.

The ray of unboundedness in the original variables is  $x_1 = 5 + t$ ,  $x_2 = 3 + 2t$ ,  $x_3 = t$ ,  $t \geq 0$ .

## Question 2

(Farkas Lemma)

We have that there exists a vector  $z \leq \mathbf{0}$  such that  $Bz - Cz = v$ . Which means that for  $x = -z$  it holds that

$$\begin{aligned} (C - B)x &= v, \\ x &\geq \mathbf{0}. \end{aligned}$$

Using Farkas lemma we then know that there can not exist any  $u \in \mathbb{R}^m$  such that

$$\begin{aligned} (C - B)^T u &\geq \mathbf{0}, \\ v^T u &< 0. \end{aligned}$$

So there can not exist any  $y \in \mathbb{R}^m$  with  $C^T y \leq B^T y$  and  $v^T y > 0$ .



**(3p) Question 3**

(KKT conditions)

- (1p)** a) Set  $f(\mathbf{x}) = -c^t\mathbf{x}$ ,  $g(\mathbf{x}) = \mathbf{x}^t\mathbf{x} - 1$ . The KKT conditions are

$$\begin{aligned}\nabla f(\mathbf{x}) + \mu\nabla g(\mathbf{x}) &= -c + 2\mu\mathbf{x}, \\ \mu g(\mathbf{x}) &= 0, \\ \mu &\geq 0.\end{aligned}$$

When  $\bar{\mathbf{x}} = c/\|c\|$ ,  $\mu = \|c\|/2$ , all the conditions are fulfilled. So  $\bar{\mathbf{x}}$  is a KKT point.

- (2p)** b) Since the objective function and the feasible set are both convex, the problem is convex. Thus KKT conditions are sufficient. Since the feasible set is convex and  $\mathbf{0}$  is an interior point, Slater CQ holds. Thus KKT conditions are necessary. To solve the KKT system, suppose  $\tilde{\mathbf{x}}$  is a KKT point. If  $g(\tilde{\mathbf{x}}) < 0$ , then  $\mu = 0$ , but  $\nabla f(\mathbf{x}) = c \neq \mathbf{0}$ , contradiction. Thus  $g(\tilde{\mathbf{x}}) = 0$ ,  $\mu > 0$ .  $\tilde{\mathbf{x}} = c/2\mu$ , plug it into  $g(\tilde{\mathbf{x}}) = 0$ , we get  $\tilde{\mathbf{x}} = c/\|c\|$ . So,  $\bar{\mathbf{x}}$  is an unique KKT point. Since KKT conditions are both necessary and sufficient,  $\bar{\mathbf{x}}$  is an unique global optimal.

**(3p) Question 4**

(Gradient projection)

*Iteration 1:* We have  $\nabla f(\mathbf{x}^0) = (-2, -3)^T$ . We need to project the point  $(0, 0)^T - (-2, -3)^T = (2, 3)^T$  on the feasible region  $X$ . We graphically see that this projection is obtained by taking the point  $(2, 2)$ . Hence,  $\mathbf{x}^1 = (2, 2)^T$ .

*Iteration 2:* We have  $\nabla f(\mathbf{x}^1) = (-2, 1)^T$ . We need to project the point  $(2, 2)^T - (-2, 1)^T = (4, 1)^T$  on the feasible region  $X$ . We graphically see that this projection is obtained by taking the point  $(3, 1)$ . Hence,  $\mathbf{x}^2 = (3, 1)^T$ .

The obtained point is neither a global nor a local minimum. This can be checked by, e.g., the KKT conditions and realizing that the point is not a stationary point.

(3p) **Question 5**

(modeling)

(1.5p) a) Definitions of additional sets

- $I := \{1, \dots, 9\}$  be the index set of rows.
- $J := \{1, \dots, 9\}$  be the index set of columns.
- $L := \{1, \dots, 9\}$  be the index set of cells.
- $K := \{1, \dots, 9\}$  be the index set of numbers.

The set of feasible solution  $S$  to the Sudoku is defined by:

$$\begin{aligned} \sum_{i \in I} x_{ijk} &= 1, & j \in J, k \in K, \\ \sum_{j \in J} x_{ijk} &= 1, & i \in I, k \in K, \\ \sum_{(i,j) \in C_l} x_{ijk} &= 1, & l \in L, k \in K, \\ \sum_{k \in K} x_{ijk} &= 1, & i \in I, j \in J, \\ x_{ijk} &= 1, & (i, j, k) \in A, \\ x_{ijk} &\in \{0, 1\}, & i \in I, j \in J, k \in K. \end{aligned}$$

(1.5p) b) Consider the objective function, to be minimized

$$f(x) := \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \bar{x}_{ijk} x_{ijk}.$$

Let  $\tilde{\mathbf{x}} \in S$  and assume that  $\tilde{\mathbf{x}} \neq \bar{\mathbf{x}}$ . Let  $\bar{k}_{ij}$  be the number assigned to tile  $(i, j)$  in solution  $\bar{\mathbf{x}}$ . Note that there exists by assumption at least one tile  $(i, j)$  such that  $\tilde{x}_{ij\bar{k}_{ij}} = 0$ . We yield that

$$f(\tilde{\mathbf{x}}) = \sum_{i \in I} \sum_{j \in J} \tilde{x}_{ij\bar{k}_{ij}} < \sum_{i \in I} \sum_{j \in J} 1 = \sum_{i \in I} \sum_{j \in J} \bar{x}_{ij\bar{k}_{ij}} \bar{x}_{ij\bar{k}_{ij}} = f(\bar{\mathbf{x}}).$$

Thus,  $\bar{\mathbf{x}}$  is not an optimal solution. □

**Question 6**

(true or false)

(1p) a) True. By Weierstrass theorem,  $f(\mathbf{y}) = \min_{\mathbf{x} \in S} \|\mathbf{y} - \mathbf{x}\|$  has an optimal solution.

Suppose the optimal solution for  $f(\mathbf{y}^1)$  is  $\mathbf{x}^1$ . For  $f(\mathbf{y}^2)$  the optimal solution is  $\mathbf{x}^2$ .

$$\begin{aligned}
 & \lambda f(\mathbf{y}^1) + (1 - \lambda)f(\mathbf{y}^2) \\
 &= \lambda \min_{\mathbf{x} \in S} \{\|\mathbf{y}^1 - \mathbf{x}\|\} + (1 - \lambda) \min_{\mathbf{x} \in S} \{\|\mathbf{y}^2 - \mathbf{x}\|\} \\
 &= \lambda \|\mathbf{y}^1 - \mathbf{x}^1\| + (1 - \lambda)\|\mathbf{y}^2 - \mathbf{x}^2\| \\
 & \quad (\text{by triangle-inequality}) \\
 & \geq \|\lambda(\mathbf{y}^1 - \mathbf{x}^1) + (1 - \lambda)(\mathbf{y}^2 - \mathbf{x}^2)\| \\
 &= \|\lambda\mathbf{y}^1 + (1 - \lambda)\mathbf{y}^2 - (\lambda\mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2)\| \\
 & \quad \text{since } S \text{ is convex, } \mathbf{x}^1 \text{ and } \mathbf{x}^2 \text{ belong to } S, \lambda\mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2 \text{ also belong to } S \\
 & \geq \min_{\mathbf{x} \in S} \{\|\lambda\mathbf{y}_1 + (1 - \lambda)\mathbf{y}_2 - \mathbf{x}\|\} \\
 &= f(\lambda\mathbf{y}_1 + (1 - \lambda)\mathbf{y}_2)
 \end{aligned}$$

Thus, the function  $f$  is convex.

- (1p) b) False. Suppose the feasible set is  $x_1^2 + x_2 \leq 0$ ,  $x_1^2 - x_2 \leq 0$ , and the objective function (to be minimized) is  $f = x_1$ . Since the only feasible point is  $(0, 0)^T$ , and the objective function is convex, the problem is convex. Thus, the KKT conditions are sufficient. But at point  $(0, 0)^T$ , the gradient cone is  $(a, 0)^T$  where  $a \in \mathbb{R}$ , and the tangent cone is  $(0, 0)^T$ , so they are not the same. Thus, the KKT conditions are not necessary.
- (1p) c) False. If no feasible solution exists, the optimal value is  $> 0$ . If feasible solutions exist, the optimal value is  $= 0$ .

### (3p) Question 7

(Lagrangian relaxation and decomposition)

- (1p) a) The Lagrangian dual function is

$$h(\mathbf{u}) = \inf \left\{ \left( 1 - \sum_{i \in \mathcal{I}} u_i \right) z + \sum_{i \in \mathcal{I}} u_i \sum_{j \in \mathcal{J}} p_{ij} x_{ij} \mid \sum_{i \in \mathcal{I}} x_{ij} = 1, j \in \mathcal{J}, x_{ij} \in \mathbb{B}, z \in \mathbb{R} \right\}$$

Since there are no constraints on  $z$  we yield that  $h(\mathbf{u}) = -\infty$  unless the coefficient  $1 - \sum_{i \in \mathcal{I}} u_i$  is zero, i.e.,  $\sum_{i \in \mathcal{I}} u_i = 1$ .

- (1.5p) b) Note that there is no constraint that connects variables from different tasks and the objective is linear. By also assuming  $\sum_{i \in \mathcal{I}} \bar{u}_i = 1$  we yield

$$h(\bar{\mathbf{u}}) = \sum_{j \in \mathcal{J}} \min \left\{ \sum_{i \in \mathcal{I}} \bar{u}_i p_{ij} x_{ij} \mid \sum_{i \in \mathcal{I}} x_{ij} = 1, x_{ij} \in \mathbb{B}, i \in \mathcal{I} \right\}$$

The constraints can be read as choose one machine for each task, hence choosing a machine with (tied) smallest objective coefficient is optimal. Hence, let  $i_j^* \in$

$\operatorname{argmin}_{i \in I} \bar{u}_i p_{ij}$ ,  $j \in J$ . The minimizer of the Lagrangian function at  $\bar{u}$  is thus  $\bar{x}_{i^*j} = 1$  for  $j \in J$  and otherwise zero. We yield

$$h(\bar{\mathbf{u}}) = \sum_{j \in J} \min_{i \in I} \bar{u}_i p_{ij}$$

- (0.5p)** c) All relaxed constraints are satisfied by choosing  $\bar{z} = \operatorname{argmax}_{i \in I} \sum_{j \in J} p_{ij} \bar{x}_{ij}$ , hence  $(\bar{\mathbf{x}}, \bar{z})$  forms a primal feasible solution.
-