Chalmers/GU
Mathematics sciences

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	20–10-29
Time:	8^{30}–13^{30}
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

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := |x_1| + |x_2|, \\ \text{subject to} \quad & x_1 - 2x_2 \geq 1, \\ & -x_1 - x_2 \leq 5. \end{aligned}$$

(1p) a) Rewrite the problem to standard form, by using the transformation $|x_i| = x_i^+ + x_i^-$ where $x_i = x_i^+ - x_i^-$ and $x_i^+, x_i^- \geq 0$, which is to be motivated in c). Then verify that using x_1^- and x_2^- as basic variables results in a basic feasible solution (BFS).

(1.5p) b) Solve the problem using the second phase of the simplex method. Use the BFS suggested in a) as the initial basis. Present the optimal solution in terms of the original variables.

(0.5p) c) Motivate the transformation made in a) by proving that the equality $|x_i| = x_i^+ + x_i^-$, $i = 1, 2$, holds in any BFS.

(3p) Question 2

(unconstrained optimization)

Consider the unconstrained problem to minimize the function

$$f(x_1, x_2) = x_1^2 + x_1x_2 - x_2^2 + 2x_1$$

(1p) a) Start in $\boldsymbol{x}^0 = (0, 0)^{\mathrm{T}}$ and perform two iterations with the steepest descent method using the step length $\alpha_k = 1$ in each iteration. Is the point reached an optimal solution?

(2p) b) Start in $\boldsymbol{x}^0 = (0, 0)^{\mathrm{T}}$ and perform two iterations with the Newton method using the Levenberg-Marquardt modification with $\gamma = 3$. Use step length $\alpha_k = 1$ in each iteration. Is the point reached an optimal solution?

(3p) Question 3

Let $\phi : \mathbf{X} \times \mathbf{Y} \mapsto \mathbb{R}$ be a continuous function, where $\mathbf{X} \subseteq \mathbb{R}^n$ and $\mathbf{Y} \subseteq \mathbb{R}^m$ are non-empty sets.

(1p) a) Show that the following inequality holds

$$\sup_{\mathbf{y}\in\mathbf{Y}} \inf_{\mathbf{x}\in\mathbf{X}} \phi(\mathbf{x},\mathbf{y}) \leq \inf_{\mathbf{x}\in\mathbf{X}} \sup_{\mathbf{y}\in\mathbf{Y}} \phi(\mathbf{x},\mathbf{y}).$$

(2p) b) Suppose that $\mathbf{X}$ and $\mathbf{Y}$ are nonempty, compact, and convex sets, and that the function ϕ is convex in $\mathbf{x}$ for any given $\mathbf{y}$ and concave in $\mathbf{y}$ for any given $\mathbf{x}$.

Show that the function $\rho : \mathbf{X} \mapsto \mathbb{R}$, defined by $\rho(\mathbf{x}) := \max_{\mathbf{y}\in\mathbf{Y}} \phi(\mathbf{x},\mathbf{y})$, is a convex function in $\mathbf{x}$ and that the function $\delta : \mathbf{Y} \mapsto \mathbb{R}$, defined by $\delta(\mathbf{y}) := \min_{\mathbf{x}\in\mathbf{X}} \phi(\mathbf{x},\mathbf{y})$, is a concave function in $\mathbf{y}$.

(3p) Question 4

(KKT conditions)

Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := -x_1 + x_2, \\ \text{subject to} \quad & x_1^2 + x_2 \leq 43, \\ & (x_1 - 1)^3 - x_2 \leq 0, \\ & x_1 \geq 2. \end{aligned}$$

(2p) a) State the KKT-conditions for the problem and check whether they are necessary or not, and whether they are sufficient or not.

(1p) b) Find all KKT-points. For each of the KKT points, state whether it is optimal. Motivate!

(3p) Question 5

(Modelling)

Consider a network flow problem on a set of nodes N and a set of edges $E \subset N \times N$. Some nodes are source nodes $S \subset N$ in which a fluid enters the system. This fluid contains some pollutant. Let $\bar{p}_i$ denote the known portion of the pollutant in the fluid leaving node $i \in S$. Fluid can be purchased in the source nodes $i \in S$ at a cost of c_i SEK per unit of fluid. Then there are some intermediate nodes $I \subset N \setminus S$ that are pools in which the incoming fluids are mixed to a homogeneous state. Finally, there are some terminal nodes $T = N \setminus (I \cup S)$, which has a quantity demand $d_i \geq 0$ of fluid to be delivered and quality requirement represented by an upper bound $\bar{p}_i$ on the allowed portion of the pollutant. You can assume that the source nodes S can not have incoming flow from other nodes and that the terminal nodes T can not have outgoing flow to other nodes.

Construct a non-linear model minimizing the total purchase cost while satisfying the demand and the desired quality in all terminal nodes.

Hints: To model this you need two sets of variables: (i) let f_{ij} the denote the amount of fluid sent from a node $i \in N$ to a node $j \in N$ and (ii) let p_i denote the portion of the pollutant in the flow leaving node $i \in N$. The amount of the pollutant in the flow from a node i to a node j can thus be computed by the expression $f_{ij}p_i$.

All intermediate nodes have equal amounts of total incoming and outgoing fluid. And similarly, the total amount (not portion) of the pollutant entering and leaving these nodes must also be equal.

Question 6

(true or false)

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

(1p) a) Consider a primal-dual pair of linear programs.

Claim: If the dual problem is infeasible then the primal problem is unbounded.

(1p) b) Consider the problem to $\min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x})$ where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a twice differentiable function (i.e., $f \in C^2$).

Claim: If $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ and $\nabla^2 f(\boldsymbol{x}^*) \succeq 0$ then $\boldsymbol{x}^*$ is a local minimum of f.

(1p) c) Consider the Frank-Wolfe method used for minimizing a non-linear function over a polyhedron.

Claim: In each iteration of the algorithm a linear program needs to be solved in order to find the search direction.

(3p) Question 7

(Lagrangian duality)

Consider the problem (P) to

$$\begin{aligned} \text{minimize} \quad & x_1^2 + 2x_2^2, \\ \text{subject to} \quad & x_1 + x_2 \geq 2, \\ & x_1, x_2 \leq 2. \end{aligned}$$

Lagrangian relax the constraint $x_1 + x_2 \geq 2$. State and evaluate the Lagrangian dual function q at $\mu = 0$ and $\mu = 6$ and provide the corresponding lower bounds on the optimal objective value to the problem (P).

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 20–10–29
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)

(1p) a) The problem on standard form is:

$$\begin{aligned} \text{minimize} \quad & x_1^+ + x_1^- + x_2^+ + x_2^- && (1)\\ \text{subject to} \quad & x_1^+ - x_1^- - 2x_2^+ + 2x_2^- - s_1 = 1, && (2)\\ & -x_1^+ + x_1^- - x_2^+ + x_2^- + s_2 = 5, && (3)\\ & x_1^+, \ x_1^-, \ x_2^+, \ x_2^-, \ s_1, s_2 \geq 0 && (4) \end{aligned}$$

Using $x_B = (x_1^-, x_2^-)$, we get

$$x_B = B^{-1}b = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \geq 0$$

hence x_B is a BFS.

(1.5p) b) First iteration: we have $x_B = (x_1^-, x_2^-)$, $x_N = (x_1^+, x_2^+, s_1, s_2)$, $B = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$, $B^{-1} = \frac{1}{3} \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$, $N = \begin{bmatrix} 1 & -2 & -1 & 0 \\ -1 & -1 & 0 & 1 \end{bmatrix}$, $c_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $c_N^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}$, $B^{-1}b = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

Checking optimality:

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

Not optimal, minimum reduce costs indicate s_2 enter the basis.

Minimum ratio test:

$$\underset{i \in (B^{-1}N_4)_i > 0}{\operatorname{argmin}} \frac{(B^{-1}b)_i}{(B^{-1}N_4)_i} = \operatorname{argmin}\{\frac{3}{2/3}, \frac{2}{1/3}\} = \operatorname{argmin}\{\frac{9}{2}, 6\}$$

hence, x_1^- leaves the basis.

Second iteration: we have $x_B = (x_2^-, s_2)$, $x_N = (x_1^+, x_1^-, x_2^+, s_1)$, $B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$, $B^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$, $N = \begin{bmatrix} 1 & -1 & -2 & -1 \\ -1 & 1 & -1 & 0 \end{bmatrix}$, $c_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $c_N^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 & 0 \end{bmatrix}$, $B^{-1}b = \frac{1}{2} \begin{bmatrix} 1 \\ 9 \end{bmatrix}$.

Checking optimality:

$$\bar{c}_N^{\mathrm{T}} = c_N^{\mathrm{T}} - c_B^{\mathrm{T}} B^{-1} N = \begin{bmatrix} 1 & 1 & 1 & 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & -2 & -1 \\ -1 & 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} & 2 & \frac{1}{2} \end{bmatrix} \geq 0$$

The current basis is optimal.

The solution in the original variables are $x_1 = 0, x_2 = -\frac{1}{2}$.

(0.5p) c) Note that the columns of x_1^+, x_1^- are linearly dependent, hence, by definition of basis they both cannot be non-zero in a BFS. Thus in every BFS one of them is non-zero and the equality hold.

Question 2

(unconstrained optimization)

We have that

$$\nabla f(\boldsymbol{x}) = (2x_1 + x_2 + 2, x_1 - 2x_2), \quad \nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}, \quad \nabla^2 f(\boldsymbol{x}) + \gamma I = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}$$

(1.5p) a) At $\boldsymbol{x}^0 = (0,0)$ the search direction is $-\nabla f(\boldsymbol{x}^0) = (-2, 0)$. So $\boldsymbol{x}^1 = (-2, 0)$.
At $\boldsymbol{x}^1 = (-2, 0)$ the search direction is $-\nabla f(\boldsymbol{x}^1) = (2, 2)$. So $\boldsymbol{x}^2 = (0, 2)$.
$\boldsymbol{x}^2 = (0, 2)$ is not an optimal solution since $\nabla f(\boldsymbol{x}^2) \neq \mathbf{0}$

(1.5p) b) At $\boldsymbol{x}^0 = (0,0)$ the search direction is $-(\nabla^2 f(\boldsymbol{x}^0) + \gamma I)^{-1} \nabla f(\boldsymbol{x}^0) = (-1/2, 1/2)$. So $\boldsymbol{x}^1 = (-1/2, 1/2)$.
At $\boldsymbol{x}^1 = (-1/2, 1/2)$ the search direction is $-(\nabla^2 f(\boldsymbol{x}^1) + \gamma I)^{-1} \nabla f(\boldsymbol{x}^1) = (3/2, -9/4)$. So $\boldsymbol{x}^2 = (1/4, -7/4)$.
$\boldsymbol{x}^2 = (1/4, -7/4)$ is not an optimal solution since $\nabla f(\boldsymbol{x}^2) \neq \mathbf{0}$

Question 3

(1p) a) Define $f(y) = \inf_{\mathbf{x} \in \mathbf{X}} \phi(\mathbf{x}, \mathbf{y})$, then it holds that $f(y) = \inf_{\mathbf{x} \in \mathbf{X}} \phi(\mathbf{x}, \mathbf{y}) \leq \phi(\mathbf{x}, \mathbf{y})$. Therefore, $\sup_{\mathbf{y} \in \mathbf{Y}} f(y) \leq \sup_{\mathbf{y} \in \mathbf{Y}} \phi(\mathbf{x}, \mathbf{y})$ for any $\mathbf{x}$.
So, $\sup_{\mathbf{y} \in \mathbf{Y}} f(y) \leq \inf_{\mathbf{x} \in \mathbf{X}} \sup_{\mathbf{y} \in \mathbf{Y}} \phi(\mathbf{x}, \mathbf{y})$. Which means:

$$\sup_{\mathbf{y} \in \mathbf{Y}} \inf_{\mathbf{x} \in \mathbf{X}} \phi(\mathbf{x}, \mathbf{y}) \leq \inf_{\mathbf{x} \in \mathbf{X}} \sup_{\mathbf{y} \in \mathbf{Y}} \phi(\mathbf{x}, \mathbf{y})$$

(2p) b)

$$\begin{aligned}
&\rho((1-\alpha)\mathbf{x_1} + \alpha \mathbf{x_2}) \\
=&\max_{\mathbf{y} \in \mathbf{Y}} \phi((1-\alpha)\mathbf{x_1} + \alpha \mathbf{x_2}, \mathbf{y}) \\
=&(\text{suppose the optimal } \mathbf{y} \text{ for this optimization problem is } \mathbf{y_1}) \\
=&\phi((1-\alpha)\mathbf{x_1} + \alpha \mathbf{x_2}, \mathbf{y_1}) \\
=&(\text{the function } \phi \text{ is convex in } \mathbf{x} \text{ for any given } \mathbf{y}) \\
\leq&(1-\alpha)\phi(\mathbf{x_1}, \mathbf{y_1}) + \alpha \phi(\mathbf{x_2}, \mathbf{y_1}) \\
\leq&(1-\alpha) \max_{\mathbf{y} \in \mathbf{Y}} \phi(\mathbf{x_1}, \mathbf{y}) + \alpha \max_{\mathbf{y} \in \mathbf{Y}} \phi(\mathbf{x_2}, \mathbf{y}) \\
=&(1-\alpha)\rho(\mathbf{x_1}) + \alpha \rho(\mathbf{x_2})
\end{aligned}$$

By definition of convexity $\rho(\mathbf{x})$ is convex.

To show the function $\min_{\mathbf{x} \in \mathbf{X}} \phi(\mathbf{x}, \mathbf{y})$ is a concave function in $\mathbf{y}$ is the same as shown $-\min_{\mathbf{x} \in \mathbf{X}} \phi(\mathbf{x}, \mathbf{y})$ is a convex function in $\mathbf{y}$, which is the same as shown $\max_{\mathbf{x} \in \mathbf{X}} -\phi(\mathbf{x}, \mathbf{y})$ is a convex function in $\mathbf{y}$. We know the function ϕ is concave in $\mathbf{y}$ for any given $\mathbf{x}$, so the function $-\phi$ is convex in $\mathbf{y}$ for any given $\mathbf{x}$. Then the rest of the prove is as before.

(3p) ## Question 4

(KKT conditions)

(2p) a) The KKT conditions are

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

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

$$\mu_i \geq 0. \qquad i=1,2,3$$

For necessity we check LICQ. For interior points, since there is no active constraints, so the gradients of the active constraints are linearly independent. For the points on the boundary, but not extreme points, since there is only one active constraint, so the gradients of the active constraints are linearly independent. Now we check the extreme points. There are three extreme points: $(2,1)^T$, $(2,39)^T$, $(4,27)^T$.

For the point $(2,1)^T$, the gradients of the active constraints are $(3,-1)^T$ and $(-1,0)$. They are linearly independent.

For the point $(2,39)^T$, the gradients of the active constraints are $(2,1)^T$ and $(-1,0)$. They are linearly independent.

For the point $(4,27)^T$, the gradients of the active constraints are $(8,1)^T$ and $(27,-1)$. They are linearly independent.

So, LICQ holds at all feasible points, which means KKT conditions are necessary.

For sufficiency, the objective function is obviously convex. $f=x_1^2+x_2$ is convex, by level set theorem, set $\{x_1^2+x_2\leq 43\}$ is convex. The eigenvalues of hessian of $\bar{f}=(x_1-1)^3-x_2$ are $6(x_1-1)$ and 0. So when $x_1\geq 2$, the function $\bar{f}$ is convex. So the set $\{(x_1-1)^3-x_2\leq 0,\ x_1\geq 2\}$ is convex. The intersection of convex sets are convex, so the feasible set is convex. Thus, the problem is convex. Which means KKT conditions are sufficient.

(1p) b) Look at the first KKT condition, we can see μ_2 must be positive. If g_2 is the only active constraint, then $x_1<2$, which is not feasible. If g_1 and g_2 are active, it corresponds to the point $(4,27)^T$.

$$\begin{pmatrix}-1\\1\end{pmatrix}+\mu_1\begin{pmatrix}8\\1\end{pmatrix}+\mu_2\begin{pmatrix}27\\-1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},$$

Solve this we get $\mu_1=-\frac{26}{35}$, $\mu_2=\frac{9}{35}$. Since $\mu_1<0$, so it is not a KKT point.

If g_2 and g_3 are active, it corresponds to the point $(2,1)^T$.

$$\begin{pmatrix}-1\\1\end{pmatrix}+\mu_2\begin{pmatrix}3\\-1\end{pmatrix}+\mu_3\begin{pmatrix}-1\\0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},$$

Solve this we get $\mu_2=1$, $\mu_3=2$. So it is a KKT point.

Since the KKT conditions are sufficient for optimality, so $(2,1)^T$ is the optimal point and the optimal value is -1.

(3p) **Question 5**

(modeling)

Additional sets:

- A set of arcs with no incoming arcs to the source nodes and no outgoing from the terminal nodes.
- $\delta^+(i)$ be the set of nodes $j \in N$ such that $(i,j) \in A$.
- $\delta^-(i)$ be the set of nodes $j \in N$ such that $(j,i) \in A$.

Variables:

- f_{ij} denote the units of flow sent from node $i \in N$ to node $j \in N$, where $(i,j) \in A$.
- p_i denote the portion of the pollutant in the flow leaving node $i \in N$.

Additional parameters:

- $\bar{p}_i$ be the known level of the pollutant leaving the source nodes $i \in S$.

$$
\begin{array}{llrr}
\text{minimize} & \displaystyle\sum_{i\in S} c_i \sum_{j\in\delta^+(i)} f_{ij} & & (1)\\
\text{subject to} & \displaystyle\sum_{j\in\delta^+(i)} f_{ij} - \sum_{j\in\delta^-(i)} f_{ji} = 0, & i \in I, & (2)\\
& \displaystyle\sum_{j\in\delta^-(i)} f_{ji} \geq d_i, & i \in T, & (3)\\
& \displaystyle\sum_{j\in\delta^-(i)} p_j f_{ji} - p_i \sum_{j\in\delta^-(i)} f_{ji} = 0, & i \in I \cup T, & (4)\\
& p_i = \bar{p}_i, & i \in S, & (5)\\
& p_i \leq \bar{p}_i, & i \in T, & (6)\\
& f_{ij} \geq 0, & (i,j) \in A. & (7)
\end{array}
$$

(2) and (3) are the flow balance equations for the fluid, (4) is the flow balance equations for the pollutants, and (5), (6), (7) are the constraints on the pollutants and on the flow.

Question 6

(true or false)

(1p) a) False. The primal problem might also be infeasible.

(1p) b) False. Counter-example is $f(x) = x^3$ and the point $x^* = 0$.

(1p) c) True. In order to find the search direction one needs to solve the problem $\min_{\boldsymbol{x} \in P} \nabla f(\boldsymbol{x}^k)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^k)$ where P is the polyhedron and $\boldsymbol{x}^k$ is the current iterate. And this is a linear program.

(3p) Question 7

(Lagrangian duality)

The Lagrangian dual function is

$$\begin{aligned} q(\mu) &= \min_{x1, x_2 \leq 2} x_1^2 + 2x_2^2 + \mu(2 - x_1 - x_2) \\ &= 2\mu + \min_{x_1 \leq 2} \left(x_1^2 - \mu x_1\right) + \min_{x_2 \leq 2} \left(2x_2^2 - \mu x_2\right). \end{aligned}$$

At $\mu = 0$ the two inner optimization problems have solutions $x_1 = 0$ and $x_2 = 0$. So $q(0) = 0$.

At $\mu = 6$ the two inner optimization problems have solutions $x_1 = 2$ and $x_2 = 1.5$ so $q(6) = -0.5$
