Chalmers/GU
Mathematics sciences

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	22–10–27
Time:	8^{30}–13^{30}
Aids:	Text memory-less calculator, English-Swedish dictionary
Number of questions:	7; a passed question requires 2 points of 3.
	Questions are *not* numbered by difficulty.
	To pass requires 10 points and three passed questions.
Examiner:	Axel Ringh (073 708 23 73 and/or 031 772 12 34)

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

(Lagrangian duality)

Consider the problem

$$\begin{aligned} \text{minimize} \quad & 2x_1^2 - 3x_2, \\ \text{subject to} \quad & x_2 \le x_1^2, \\ & -1 \le x_1 \le 1, \\ & 0 \le x_2 \le 1. \end{aligned}$$

(2p) a) Lagrangian relax the first constraint with a multiplier $\mu \ge 0$, i.e., the constraint $x_2 \le x_1^2$, and derive the dual function $q(\mu)$.

Hint: For $X := \{x \in \mathbb{R}^2 \mid -1 \le x_1 \le 1,\ 0 \le x_2 \le 1\}$, an explicit solution to $\min_{x \in X} \mathscr{L}(x, \mu)$ can be found for each μ in this case.

(1p) b) Compute a subgradient to $q(\mu)$ in the point $\mu = 1$.

Question 2

(theory question - sufficiency of KKT conditions for convex problems)

Consider the problem

$$\text{(P)} \qquad \begin{cases} \text{minimize} & f(x), \\ \text{subject to} & g_i(x) \le 0, \quad i = 1, \dots, m, \\ & h_j(x) = 0, \quad j = 1, \dots, \ell, \\ & x \in \mathbb{R}^n. \end{cases}$$

(1p) a) State the Karush-Kuhn-Tucker (KKT) conditions for problem (P).

(2p) b) A point x is called a KKT point to (P) if there is a solution to the KKT conditions corresponding to (P) in the point x. Prove the following theorem:

THEOREM: *Consider the problem* (P), *and let*

$$S = \{x \in \mathbb{R}^n \mid g_i(x) \le 0,\ \textit{for } i = 1, \dots, m,\ \textit{and } h_j(x) = 0,\ \textit{for } j = 1, \dots, \ell\}.$$

Assume that the functions f and $g_i(x)$, for $i = 1, \dots, m$, are all convex, and that $h_j(x)$, for $j = 1, \dots, \ell$, are all affine. If $x^* \in S$, then

$$x^* \textit{ is a KKT point to } \text{(P)} \quad \Longrightarrow \quad x^* \textit{ is a global minimizer to } \text{(P)}.$$

(3p) Question 3

(the simplex method)

Consider the following linear program:

$$\begin{array}{rll} \text{minimize} & -x_1 - 2x_2 + 4x_3 & \\ \text{subject to} & 2x_1 + 3x_3 & \leq 7, \\ & x_1 + x_2 - 3x_3 & \leq 7, \\ & x_1,\ x_2,\ x_3 & \geq 0. \end{array}$$

Solve it using the simplex method (phase II), and start with x_1 and x_2 as basic variables.

Hint: You may find the following identity useful:

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

(3p) Question 4

(modelling)

A friend of yours, we can call the person AR, has realized that the person has too many things at home. On day, AR therefore decides to sell some of the things at a local flea market. To this end, AR first makes a list $I = \{1, 2, \ldots, n\}$ of things that are up for sale. For each object $i \in I$, AR sets the price c_i. However, for certain items the price will depend on what other items are also brought by AR to be sold at the flea market. For example: AR has two copies of the same book, and one is in better condition than the other. If both are brought to the flea market, the book in worse condition will have to be sold at a reduced price. On the other hand, AR has two candlesticks of the same sort, and selling them together can motivate a higher, premium, price compared to if only one is brought and sold. This type of price adjustments, i.e., the influence an item $i \in I$ has on price of another item $j \in I$, are recorded as p_{ij}, for $i, j \in I$. In particular, $p_{ii} = 0$ for all $i \in I$. Also note that one might have $p_{ij} \neq p_{ji}$, as in the example with the two books: the book $i \in I$ in good condition lowers the price of the book $j \in I$ in bad condition (i.e., $p_{ij} < 0$), but the book in bad condition does not affect the price of the book in good condition (i.e., $p_{ji} = 0$).

AR will bring the items to the flea market in a backpack. Each item $i \in I$ has a certain weight, and AR can carry at most a certain maximum total weight. Moreover, in order for the table at the flea market to look appealing, it should neither be crowded with things nor should it be half-empty. Therefore, AR wants to bring at least a certain number of items, but also at most a certain number of items. Help AR decide which items to bring to the flea market in order to maximize the profit (assuming that all items brought are indeed sold). Introduce appropriate constants (in addition to c_i for $i \in I$ and p_{ij} for $i, j \in I$), introduce appropriate variables, and formulate the decision problem as an integer linear program.

Question 5

(interior penalty method)

Consider the problem

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

(0.5p) a) Prove that the point

$$x^* = \begin{bmatrix} -\frac{8}{5} \\ -\frac{1}{5} \end{bmatrix}$$

is globally optimal to the problem.

(1.5p) b) Consider solving the problem using an interior penalty method. The penalty problem is to solve

$$\min_{x \in \mathbb{R}^2} f(x) + \nu \hat{\chi}(x),$$

where $\hat{\chi}(x) = \phi(g(x))$. Use the logarithmic barrier function

$$\phi(s) = -\log(-s),$$

and compute the optimal solution to the penalty problem, i.e., compute the optimal $x^*(\nu)$ as a function of ν. Show that $x^*(\nu) \to x^*$ when $\nu \to 0$.

(1p) c) For each ν, an estimate of the multiplier in the KKT system is given by

$$\mu(\nu) := \nu \phi'(g(x^*(\nu))),$$

where ϕ' denotes the derivative of ϕ. Show that $\mu(\nu) \to \mu^*$ when $\nu \to 0$, where μ^* is the multiplier in the KKT system corresponding to x^*.

Hint: You may find the following result useful (L'Hôpital's rule): Let $h_1, h_2 : \mathbb{R} \to \mathbb{R}$ be continuously differentiable, and let $h_1'(x)$ and $h_2'(x)$ denote the corresponding derivatives. If $\lim_{x \to c} h_1(x) = \lim_{x \to c} h_2(x) = 0$, and if $\lim_{x \to c} h_2'(x) \neq 0$, then

$$\lim_{x \to c} \frac{h_1(x)}{h_2(x)} = \lim_{x \to c} \frac{h_1'(x)}{h_2'(x)},$$

given that the latter limit exists.

Question 6

(True or False)

The below three claims should be assessed. For each claim: state whether it is true or false. Provide an answer together with a short but complete motivation.

(1p) a) Let

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 0 & -5 & -1 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

Claim: The polyhedron $S := \{x \in \mathbb{R}^3 \mid Ax \leq b, x \geq 0\}$ is nonempty.

(1p) b) Let $f : \mathbb{R}^2 \to \mathbb{R}$, $g_1 : \mathbb{R}^2 \to \mathbb{R}$, and $g_2 : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$\begin{aligned} f(x) &= x_1^2 + x_2^2, \\ g_1(x) &= (x_1 - 2)^2 + x_2^2 - 1, \\ g_2(x) &= x_1 - \frac{1}{2}\cos(2\pi x_2) - 2, \end{aligned}$$

respectively.

Claim: The point $x^* = [1, 0]^T$ is a globally optimal solution to

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & f(x) \\ \text{subject to} \quad & g_1(x) \leq 0 \\ & g_2(x) \leq 0. \end{aligned}$$

(1p) c) Consider $f : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$f(x) = (1 - x_1 x_2)^2 + x_1^2 + (x_1^2 + x_3)^2.$$

Claim: The optimization problem

$$\begin{aligned} \inf \quad & f(x) \\ \text{subject to} \quad & x \in \mathbb{R}^3 \end{aligned}$$

attains a global minimum.

(3p) Question 7

(The gradient projection method)

The gradient projection method can be used to solve problems of the form

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & x \in S. \end{aligned}$$

Assume that S is convex. Given an iterate x^k, the method computes a new iterate as

$$x^{k+1} = \text{Proj}_S(x^k - \alpha_k \nabla f(x^k)),$$

where

$$\text{Proj}_S(y) := \underset{x \in S}{\arg\min} \, \|x - y\|.$$

Consider $f(x) = (x_1 + 1)^2 + (x_2 + 1)^4$ and $S = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \leq 2,\ x_1 \geq 0,\ x_2 \geq 0\}$. Start in the point $x^0 = [1, 1]^T$ and perform two steps in the gradient projection algorithm using the constant step length $\alpha_k = 1/4$ for all k. The projection step in the algorithm can be solved graphically, but a clear motivation must be given. Is the obtained point x^2 locally and/or globally optimal? Motivate your answer.

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 22–10–27
Examiner: Axel Ringh

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

Question 1

(Lagrangian duality)

(2p) a) The dual function $q(\mu) = \inf_{x \in X} \mathscr{L}(x, \mu)$ is given by

$$q(\mu) = \begin{array}{rl} \min & (2-\mu)x_1^2 - (3-\mu)x_2^2 \\ \text{subject to} & -1 \leq x_1 \leq 1 \\ & 0 \leq x_2 \leq 1 \end{array} = \begin{cases} -3+\mu & \text{if } \mu \leq 2, \\ -1 & \text{if } 2 \leq \mu \leq 3, \\ 2-\mu & \text{if } \mu \geq 3, \end{cases}$$

where minimizers of the Lagrangian are attained in $x = [0, 1]^T$ for $\mu \leq 2$, in $x = [\pm 1, 1]^T$ for $2 \leq \mu \leq 3$, and in $x = [\pm 1, 0]^T$ for $\mu \geq 3$.

Note: for $\mu = 2$, all minimizers of the Lagrangian are given by $x = [a, 1]^T$ for $a \in [-1, 1]$, and for $\mu = 3$ they are given by $x = [\pm 1, b]^T$ for $b \in [0, 1]$.

(1p) b) The function is differentiable in the point, so the only subgradient is $\nabla q(1) = 1$.

Question 2

(theory question - sufficiency of KKT conditions for convex problems)

(1p) a) See equation (5.17) in the book.

(2p) b) See Theorem 5.49 in the book.

(3p) Question 3

(the simplex method)

Transforming the problem to standard form gives

$$\begin{aligned} \text{minimize} \quad & -x_1 - 2x_2 + 4x_3 \\ \text{subject to} \quad & 2x_1 + 3x_3 + s_1 = 7, \\ & x_1 + x_2 - 3x_3 + s_2 = 7, \\ & x_1,\, x_2,\, x_3,\, s_1,\, s_2 \geq 0. \end{aligned}$$

Iteration 1:
With $x_B = [x_1,\, x_2]^T$ and $x_N = [x_3,\, s_1,\, s_2]^T$,

$$B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} 3 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}, \quad c_B^T = \begin{bmatrix} -1 & -2 \end{bmatrix}, \quad c_N^T = \begin{bmatrix} 4 & 0 & 0 \end{bmatrix}.$$

$x_B = B^{-1}b = [7/2,\, 7/2]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [-7/2,\, 1/2,\, -2]$, and hence $(x_N)_1 = x_3$ enters the basis. $B^T N_1 = [3/2,\, -9/2]^T$, and since only the first component is positive we have that $(x_B)_1 = x_1$ leaves the basis.

Iteration 2:
With $x_B = [x_2,\, x_3]^T$ and $x_N = [x_1,\, s_1,\, s_2]^T$,

$$B = \begin{bmatrix} 0 & 3 \\ 1 & -3 \end{bmatrix}, \quad N = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad c_B^T = \begin{bmatrix} -2 & -4 \end{bmatrix}, \quad c_N^T = \begin{bmatrix} -1 & 0 & 0 \end{bmatrix}.$$

$x_B = B^{-1}b = [14,\, 7/3]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [7/3,\, 2/3,\, 2] \geq 0$. Hence the point $x^* = [0,\, 14,\, 7/3]^T$ is optimal.

(3p) Question 4

(modelling)

For each $i \in I$, introduce

$$x_i = \begin{cases} 1 & \text{if items } i \text{ is brought to the flea market} \\ 0 & \text{else.} \end{cases}$$

For each $(i, j) \in I \times I$, introduce

$$z_{ij} = \begin{cases} 1 & \text{if item } i \text{ and } j \text{ are both brought to the flea market} \\ 0 & \text{else.} \end{cases}$$

An integer linear program for maximizing the sell price can be formulated as

$$\begin{aligned}
\max \quad & \sum_{i=1}^{n} c_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} z_{ij} \\
\text{subject to} \quad & \sum_{i=1}^{n} m_i x_i \leq M && \text{(total weight constraint)} \\
& \sum_{i=1}^{n} x_i \geq a, && \text{(not too few items)} \\
& \sum_{i=1}^{n} x_i \leq b, && \text{(not too many items)} \\
& z_{ij} \leq \frac{1}{2}(x_i + x_j), \quad \forall i, j \in I, && (1) \\
& z_{ij} \geq -1 + x_i + x_j \quad \forall i, j \in I, && (2) \\
& x_i \in \{0, 1\}, \quad \forall i \in I, \\
& z_{ij} \in \{0, 1\}, \quad \forall i, j \in I.
\end{aligned}$$

The constraints (1) and (2), together with $z_{ij} \in \{0, 1\}$, make sure that $z_{ij} = 0$ if at most one of the items $i \in I$ and $j \in I$ are brought, and $z_{ij} = 1$ if both are brought.

Question 5

(interior penalty method)

The given problem is

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

(0.5p) a) The problem is convex (need to motivate!) and since all constraints are affine Abadie's CQ holds. Therefore, KKT is both necessary and sufficient for global optimality. Moreover, $g(x^*) = 0$, so the constraint is active. The KKT system in the point x^* is thus

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 * \left(-\frac{8}{5}\right) \\ 2 * \left(-\frac{1}{5}\right) + 2 \end{bmatrix} + \mu \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{16}{5} \\ \frac{8}{5} \end{bmatrix} + \mu \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

which has a solution $\mu^* = 8/5$. Hence the point is globally optimal.

(1.5p) b) The barrier transformed problem takes the form

$$\min x_1^2 + (x_2+1)^2 - \nu \log(-2x_1 + x_2 - 3).$$

This is a convex problem on $X := \{x \in \mathbb{R}^2 \mid g(x) < 0\}$, since $f(x) = x_1^2 + (x_2+1)^2$ is convex, and since $g(x) = 2x_1 - x_2 + 3$ is convex and $\phi(s)$ is convex and non-decreasing on X (see Proposition 3.44).

The global optimal solution to the barrier transformed problem is thus given by

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2x_1 + 2\nu \frac{1}{-2x_1+x_2-3} \\ 2x_2 + 2 - \nu \frac{1}{-2x_1+x_2-3} \end{bmatrix}.$$

From this, it follows that $2x_1 + 2(2x_2 + 2) = 0$, and solving this for x_2 gives

$$x_2 = -\frac{1}{2}x_1 - 1.$$

Substituting this back into the first equation above and rearranging terms gives

$$\frac{5}{2}x_1^2 + 4x_1 - \nu = 0 \quad \Longrightarrow \quad x_1(\mu) = -\frac{4}{5} \pm \sqrt{\frac{16}{25} + \frac{2}{5}\nu}.$$

This in turn gives

$$x_2(\nu) = -\frac{1}{2}x_1(\nu) - 1 = -\frac{3}{5} \mp \frac{1}{2}\sqrt{\frac{16}{25} + \frac{2}{5}\nu}.$$

To see which sign to use we look at which point is strictly feasible:

$$0 \geq 2x_1(\mu) - x_2(\mu) + 3 = \ldots = 2 \pm \frac{5}{3}\sqrt{\frac{16}{25} + \frac{2}{5}\nu},$$

and hence we need to choose the negative root for the first cooridnate and thus the positive root for the second coordinate, i.e.,

$$x_1^*(\mu) = -\frac{4}{5} - \sqrt{\frac{16}{25} + \frac{2}{5}\nu},$$
$$x_2^*(\mu) = -\frac{3}{5} + \frac{1}{2}\sqrt{\frac{16}{25} + \frac{2}{5}\nu}.$$

Now

$$\lim_{\nu \to 0} x^*(\nu) = \lim_{\nu \to 0} \begin{bmatrix} -\frac{4}{5} - \sqrt{\frac{16}{25} + \frac{2}{5}\nu}, \\ -\frac{3}{5} + \frac{1}{2}\sqrt{\frac{16}{25} + \frac{2}{5}\nu} \end{bmatrix} = \begin{bmatrix} -\frac{4}{5} - \frac{4}{5}, \\ -\frac{3}{5} + \frac{1}{2}\frac{4}{5} \end{bmatrix} = \begin{bmatrix} -\frac{8}{5} \\ -\frac{1}{5} \end{bmatrix} = x^*.$$

(1p) c) We get that

$$\mu(\nu) := \nu\phi'(g(x(\nu))) = \frac{\nu}{-2x_1^*(\nu) + x_2^*(\nu) - 3} = \ldots = \frac{\nu}{-2 + \frac{5}{2}\sqrt{\frac{16}{25} + \frac{2}{5}\nu}} := \frac{h_1(\nu)}{h_2(\nu)}.$$

Using L'Hôpital's rule, we find that

$$\lim_{\nu \to 0} \frac{h_1(\nu)}{h_2(\nu)} = \lim_{\nu \to 0} \frac{h_1'(\nu)}{h_2'(\nu)} = \lim_{\nu \to 0} \frac{1}{\frac{5}{2}\frac{-\frac{1}{2}}{\sqrt{\frac{16}{25}+\frac{2}{5}\nu}}\frac{2}{5}} = \frac{8}{5} = \mu^*$$

Question 6

(True or False)

The below three claims should be assessed. For each claim: state whether it is true or false. Provide an answer together with a short but complete motivation.

(1p) a) True. E.g., the point $x = [0, 1, 1]^T$ can be verified to be in the polyhedron.

Another approach: introduce slack variables and consider the polyhedron $S' = \{\tilde{A}\tilde{x} = b, \tilde{x} \geq 0\}$, where

$$\tilde{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix}, \qquad A = \begin{bmatrix} -1 & 1 & 2 & 1 & 0 \\ 0 & -5 & -1 & 0 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

S is nonempty if and only if S' is. By Farkas' lemma, exactly one of the two systems

$$\begin{cases} \tilde{A}\tilde{x} = b \\ \tilde{x} \geq 0 \end{cases} \qquad \text{and} \qquad \begin{cases} \tilde{A}^T y \leq 0 \\ b^T y > 0 \end{cases}$$

has a solution. By direct calculations, it is easily concluded that $\tilde{A}^T y \leq 0$ implies that $y = [0, 0]^T$; this can also be seen if the corresponding set is drawn graphically. Hence, the second system does not have a solution. Therefore, the first system has a solution, and thus S is nonempty.

(1p) b) True. The relaxed problem obtained by removing the constraint $g_2(x) \leq 0$ is a convex problem. It is easily verified that the given point x^* is a global optimal solution to the relaxed problem. Since it is feasible to the original problem, by the relaxation theorem it must thus be a globally opitmal solution to the original problem.

(1p) c) False. f is bounded from below by 0, but does not attain it. It is a sum of three squares and thus nonnegative. However, the first and the second square terms cannot be zero at the same time, since for the second to equal zero we must have $x_1 = 0$. Nevertheless, for any $\alpha > 0$, let $x_\alpha = [1/\alpha, \alpha, -1/\alpha^2]^T$. This gives

$$f(x_\alpha) = \left(1 - \frac{1}{\alpha}\alpha\right)^2 + \frac{1}{\alpha^2} + \left(\frac{1}{\alpha^2} - \frac{1}{\alpha^2}\right)^2 = \frac{1}{\alpha^2} \to 0$$

as $\alpha \to \infty$. Hence, the infimum is 0, but no minimizer is attained.

One way to come up with the idea for the parametrization of x_α is to try to solve $\nabla f(x) = 0$. The equation has no solution, which means that no finite point can be optimal. But the second and third component of the gradient are zero in x_α.

Note: f is not weakly coercive on $\mathbb{R}^3$, and thus Weierstrass' theorem cannot be applied.

(3p) Question 7

(The gradient projection method)

We compute the gradient of f:

$$\nabla f(x) = \begin{bmatrix} 2(x_1 + 1) \\ 4(x_2 + 1)^3 \end{bmatrix}.$$

To perform the first iteration, let

$$\tilde{x}^1 = x^0 - \frac{1}{4}\nabla f(x^0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 4 \\ 32 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \end{bmatrix}.$$

The next iterate is thus $x^1 = \text{Proj}_S(\tilde{x}^1)$, which can either be solved graphically or analytically. In any case, it is found that $x^1 = [0, 0]^T$.

To perform the second iteration, let

$$\tilde{x}^2 = x^1 - \frac{1}{4}\nabla f(x^1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix}.$$

The next iterate is thus $x^2 = \text{Proj}_S(\tilde{x}^2)$, which can either be solved graphically or analytically. In any case, it is found that $x^2 = [0, 0]^T$.

Since the cost function is convex and the constraints are given by a polyhedron, the problem is convex. Moreover, the last iteration shows that the point $x^* = [0, 0]^T$ is stationary, and since the problem is convex it is thus a globally optimal solution (see Theorem 4.23 and Proposition 4.25).