Chalmers/GU
Mathematics sciences

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	24–08–20
Time:	8^{30}–13^{30}
Aids:	Chalmers approved calculator
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.
Teacher on duty:	Ann-Brith Strömberg (070 527 36 45)
Examiner:	Axel Ringh (073 708 23 73)

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

(Linear programming)

(2p) a) Consider the linear programming problem

$$\begin{aligned} \text{minimize} \quad & 7x_1 - 8x_2 + 9x_3 \\ \text{subject to} \quad & x_1 - 2x_2 + 3x_3 \leq 1 \\ & 4x_1 - 5x_2 + 6x_3 \leq 1 \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

Solve the problem using the Simplex method.

Hint: You may find the following identity useful:

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

(1p) b) If you find that an optimal solution exists, then use your calculations to decide whether it is unique or not. If the problem is unbounded, then use your calculations to specify a ray of unboundedness of the objective value.

Question 2

(LP duality)

Consider the linear programming problem

$$\begin{aligned} \text{minimize} \quad & -4x_1 - 2x_2 - 6x_3 \\ \text{subject to} \quad & 2x_1 - x_2 + 4x_3 \leq 20 \\ & x_1 + 2x_2 + 4x_3 = 60 \\ & x_1, x_3 \geq 0. \end{aligned}$$

(2p) a) State the dual problem.

(1p) b) An optimal solution to the primal problem is

$$x^* = \begin{bmatrix} 20 \\ 20 \\ 0 \end{bmatrix}.$$

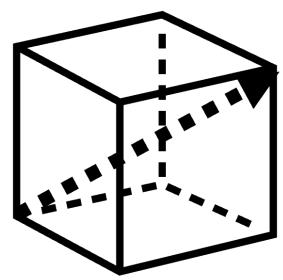
Give an optimal solution to the dual problem.

Question 3

(Modelling)

(2p) a) Consider the box shown in the figure below. It is drawn as a cube, but in this exercise we are interested in general rectangular cuboids, that is, cuboids in which all angles are right angles, and opposite faces are equal.

Formulate an optimization problem for finding the height, width, and depth of the box so that it has total volume 1, but with as short space diagonal as possible. The space diagonal is the dotted arrow in the figure.



(1p) b) If the optimization problem constructed in part a) is not a convex optimization problem, reformulate it as a convex optimization problem. Carefully motivate why the problem formulated here is convex (or why the problem you formulated in part a) is convex, if that is the case).

Hint 1: Consider the problem obtained by squaring the objective function from the problem formulated in part a). Does this new problem have the same optimal solution?

Hint 2: Consider a logarithmic transformation of the variables in the formulation in part a).

Question 4

(True or False)

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

(1p) a) Let $g_1 : \mathbb{R}^3 \to \mathbb{R}$ and $g_2 : \mathbb{R}^3 \to \mathbb{R}$ be defined by

$$g_1(x) = x_1^2 + x_2^2 + x_3^2 - 1,$$
$$g_2(x) = -(x_1 - 3)^2 - (x_2 - 3)^2 - (x_3 - 3)^2 + 1,$$

respectively, and let $S = \{x \in \mathbb{R}^3 \mid g_i(x) \leq 0 \text{ for } i = 1, 2\}$.

Claim: The set S is convex.

(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 given by

$$f(x) = 9x_1^2 + x_1 x_2 + 4x_2^2 - x_1 + 3,$$
$$g_1(x) = -(x_1 - 2)^2 - (x_2 - 2)^2 + 9,$$
$$g_2(x) = x_1^2 + x_2^2 - 4,$$

respectively, and consider the problem

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, 2. \end{cases}$$

Claim: Newton's method, as taught in the course, can be used to solve this optimization problem.

(1p) c) Consider the problem

$$\text{(P)} \qquad \begin{cases} \text{minimize} & f(x), \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \dots, m, \end{cases}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \dots, m$. Assume that the point x^* is a globally optimal solution to (P).

Claim: If the following conditions hold:

i) Abadie's constraint qualifier holds in x^*, and x^* is a KKT point,

ii) for constraint k, $g_k(x^*) = 0$ (that is, it is active), but the corresponding multiplier in the KKT system is equal to zero (that is, $\mu_k = 0$),

then x^* is also globally optimal to the problem

$$\text{(P')} \qquad \begin{cases} \text{minimize} & f(x), \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \dots, k-1, k+1, \dots, m. \end{cases}$$

(3p) Question 5

(Characterization of convexity for continuously differentiable functions)

Convex functions are an important class of functions. The following theorem is a characterization of continuously differentiable convex functions.

THEOREM: *Let $f \in \mathcal{C}^1$ on an open convex set $S \subset \mathbb{R}^n$. f is convex on S if and only if $f(y) \geq f(x) + \nabla f(x)^T(y - x)$ for all $x, y \in S$.*

Prove the above theorem. Do so using basic results from the course. If you rely on other results when performing your proof of the theorem, then those results must be stated explicitly; they may however be utilized without proof.

(3p) Question 6

(Nonlinear optimization)

Consider the optimization problem

$$\begin{cases} \underset{x \in \mathbb{R}^3}{\text{minimize}} & f(x), \\ \text{subject to} & h(x) = 0, \end{cases}$$

where

$$\begin{aligned} f(x) &= x_1^4 + x_2^2 + x_3^4, \\ h(x) &= x_1 + x_2 + x_3 - 1. \end{aligned}$$

Use results and methods from the course to find the globally optimal solution to the problem. It is ok to present the solution (and the "certificate" for global optimality) in terms of rounded numbers with 3 significant digits.

Hint 1: Do NOT try to use an iterative numerical method from the course to solve the problem. The problem can be solved analytically.

Hint 2: A so called depressed cubic equation is an equation of the form $t^3 + pt + q = 0$, where p and q are real numbers. If $\Delta := q^2/2 + p^3/27$ is greater than 0, than a real solution to the depressed cubic equation is given by

$$t = \sqrt[3]{u_1} + \sqrt[3]{u_2},$$

where $u_1 = -q/2 + \sqrt{\Delta}$ and $u_2 = -q/2 - \sqrt{\Delta}$.

Question 7

(Interior penalty methods)

Consider the problem

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

Let $\mathbb{R}_+ := \{s \in \mathbb{R} \mid s \ge 0\}$, let $\mathbb{R}_- := \{s \in \mathbb{R} \mid s \le 0\}$, and let

$$S := \{x \in \mathbb{R}^n \mid g_i(x) \le 0, \text{ for } i = 1, \dots, m\}.$$

Also consider the transformed problem

$$(\mathrm{P}_\nu) \qquad \begin{cases} \text{minimize} & f(x) + \nu \hat{\chi}_S(x), \\ \text{subject to} & x \in \mathbb{R}^n, \end{cases}$$

where

$$\hat{\chi}_S(x) := \begin{cases} \sum_{i=1}^m \phi(g_i(x)) & \text{if } g_i(x) < 0 \text{ for } i = 1, \dots, m, \\ \infty & \text{otherwise.} \end{cases}$$

(1p) a) For certain functions $\phi : \mathbb{R}_- \to \mathbb{R}_+$, which have some specific properties, we call (P_ν) an interior penalty transformation of (P). Such functions ϕ are called interior penalty functions. Which of the functions

$$\phi_1(s) = s^2, \quad \phi_2(s) = \log(-s), \quad \phi_3(s) = |s|, \quad \phi_4(s) = -\frac{1}{s-1}, \quad \phi_5(s) = \frac{1}{s^2}$$

are interior penalty functions? Here, log denotes the natural logarithm, that is, $\log(e^x) = x$.

(2p) b) A valid interior penalty function is given by $\phi(s) = -1/s$. As a specific instance of a problem of the form (P), let $n = 2$, $m = 1$, and let

$$f(x) = x_1^2 + x_2^2,$$
$$g(x) = 2x_1 - x_2 + 30.$$

Denote by $x(\nu)$ an optimal solution to the penalty transformed problem (P_ν) with penalty parameter ν. Verify that the following can be the output of an interior penalty method using this penalty function, where each step is assumed to be solved to a reasonable numerical precision.

k	ν^k	$x(\nu^k)$
0	1000	$\begin{bmatrix} -15.240 \\ 7.6201 \end{bmatrix}$
1	100	$\begin{bmatrix} -13.105 \\ 6.5525 \end{bmatrix}$
2	10	$\begin{bmatrix} -12.360 \\ 6.1799 \end{bmatrix}$
$\vdots$		

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 24–08–20
Examiner: Axel Ringh

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

(3p) Question 1

(Linear programming)

(2p) a) Transforming the problem to standard form gives

$$\begin{aligned} \text{minimize} \quad & 7x_1 - 8x_2 + 9x_3 \\ \text{subject to} \quad & x_1 - 2x_2 + 3x_3 + s_1 = 1 \\ & 4x_1 - 5x_2 + 6x_3 + s_2 = 1 \\ & x_1, x_2, x_3, s_1, s_2 \geq 0. \end{aligned}$$

From this, we see that we can start simplex (phase-II) with s_1 and s_2 as basic variables.

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

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & -2 & 3 \\ 4 & -5 & 6 \end{bmatrix}, \quad c_B^T = \begin{bmatrix} 0 & 0 \end{bmatrix}, \quad c_N^T = \begin{bmatrix} 7 & -8 & 9 \end{bmatrix}.$$

$x_B = B^{-1}b = [1,\ 1]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [7,\ -8,\ 9]$, and hence $(x_N)_2 = x_2$ enters the basis. $B^{-1}N_2 = [-2,\ -5]^T$, and hence the problem is unbounded from below.

(1p) b) A ray of unboundedness for the objective function is given by

$$x + \mu p = \begin{bmatrix} x_B \\ x_N \end{bmatrix} + \mu \begin{bmatrix} p_B \\ p_N \end{bmatrix} = \begin{bmatrix} x_B \\ x_N \end{bmatrix} + \mu \begin{bmatrix} -B^{-1}N_2 \\ e_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \mu \begin{bmatrix} 0 \\ 1 \\ 0 \\ 2 \\ 5 \end{bmatrix}$$

for $\mu \geq 0$.

(3p) Question 2

(LP duality)

(2p) a)

$$\begin{aligned} \text{maximize} \quad & 20y_1 + 60y_2 \\ \text{subject to} \quad & 2y_1 + y_2 \leq -4 \\ & -y_1 + 2y_2 = -2 \\ & 4y_1 + 4y_2 \leq -6 \\ & y_1 \leq 0 \\ & y_2 \in \mathbb{R} \end{aligned}$$

(1p) b) Since the optimal x_1 is nonzero, by complementary slackness of optimal solutions to LPs, the optimal solution to the dual problem must be tight on the first constraint. Since the second constraint in the dual is an equality constraint, the optimal solution can be found by solving the system of equations

$$\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} y_1^* \\ y_2^* \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \end{bmatrix} \qquad \Longrightarrow \qquad \begin{bmatrix} y_1^* \\ y_2^* \end{bmatrix} = \begin{bmatrix} -\frac{6}{5} \\ -\frac{8}{5} \end{bmatrix}.$$

It is easily verified that this point is also feasible for the third constraint, and that it attains the objective function value -120, which is the same as the optimal value of the primal problem.

Question 3

(Modelling)

(2p) a) Let the height, width, and depth of the box be denoted by x_1, x_2 and x_3, respectively. The volumne of the box is given by $x_1x_2x_3$ and the length of the space diagonal is given by $\sqrt{x_1^2+x_2^2+x_3^2}$. The optimization problem is therefore

$$\begin{aligned} \min_{x_1,x_2,x_3} \quad & \sqrt{x_1^2+x_2^2+x_3^2} & \\ \text{subject to} \quad & x_1x_2x_3 = 1 & \text{(volume of box is 1)} \\ & x_i > 0, \quad i=1,2,3. & \text{(lengths are positive)} \end{aligned}$$

(1p) b) Let $f(x)$ be the objective function above, and note that $f(x) \geq 0$ for all feasible points. If x^* is a globally optimal solution to the above problem, by definition $f(x^*) \leq f(x)$ for all feasible x. But this means that $f(x^*)^2 \leq f(x)^2$ for all feasible x, and hence we can equivalently consider the problem

$$\begin{aligned} \min_{x_1,x_2,x_3} \quad & x_1^2+x_2^2+x_3^2 & \\ \text{subject to} \quad & x_1x_2x_3 = 1 & \text{(volume of box is 1)} \\ & x_i > 0, \quad i=1,2,3. & \text{(lengths are positive)} \end{aligned}$$

Next, let $y_i = \log(x_i)$, for $i = 1,2,3$, where log denotes the natural logarithm. Then $x_i^2 = e^{2y_i}$, for $i = 1,2,3$, and we can rewrite the volume constraint by noting that

$$y_1+y_2+y_3 = \log(x_1)+\log(x_2)+\log(x_3) = \log(x_1x_2x_3) = \log(1) = 0.$$

Since log is bijective on $\{z \in \mathbb{R} \mid z > 0\}$, this gives the equivalent problem

$$\begin{aligned} \min_{y_1,y_2,y_3} \quad & e^{2y_1}+e^{2y_2}+e^{2y_3} \\ \text{subject to} \quad & y_1+y_2+y_3 = 0, \end{aligned}$$

which is convex (motivate!).

Question 4

(True or False)

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

(1p) a) True. For all x such that $g_1(x) \leq 0$, we have that $|x_1| \leq 1$, $|x_2| \leq 1$, and $|x_3| \leq 1$. For all such points, it is easily verified that $g_2(x) < 0$. Hence $S = \{x \in \mathbb{R}^3 \mid g_1(x) \leq 0\}$, and the latter is a convex set (why?).

(Geometric intuition: $S = \{x \in \mathbb{R}^3 \mid g_1(x) \leq 0\} \cap \{x \in \mathbb{R}^3 \mid g_2(x) \leq 0\}$. The first set is all points inside the unit sphere, and the second set is all points outside of a sphere of radius 1 centered in the point $[3, 3, 3]^T$. All points in the first set are also in the second set, and thus $S = \{x \in \mathbb{R}^3 \mid g_1(x) \leq 0\}$.)

(1p) b) False. Newton's method, as taught in this course, are used to solve unconstrained optimization problems.

(1p) c) False. $f(x) = x^3$, $g_1(x) = -x$, $g_2(x) = x^2 - 1$, and $x^* = 0$ is a counterexample. To this end, first note that $x^* = 0$ is the unique globally optimal solution (why?). Furthermore, Abadie's constraint qualifier holds in x^* (motivate!). Finally, $\nabla f(x^*) = 0$ and hence x^* is a KKT point with $\mu_1 = \mu_2 = 0$ (note that $g_2(x^*) < 0$). Nevertheless, removing the constraint g_1, the problem to minimize $f(x)$ subject to $g_2(x) \leq 0$ has the globally optimal solution $\hat{x} = -1 \neq x^*$.

(3p) Question 5

(Characterization of convexity for continuously differentiable functions)

See Theorem 3.48 a) in the course book.

(3p) Question 6

(Nonlinear optimization)

The cost function f is convex (verify!) and the constraint h is affine, which means that the problem is convex. Next, note that Slater CQ holds (no inequality constraints, and affine equality constraints). Together with convexity this means that a point is globally optimal if and only if it is a KKT point (Corollary 5.51).

A globally optimal solution (if one exists) is given by a point x such that

$$\nabla f(x) + \lambda \nabla h(x) = \begin{bmatrix} 4x_1^3 \\ 2x_2 \\ 4x_3^3 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$
$$x_1 + x_2 + x_3 = 1,$$

for some $\lambda \in \mathbb{R}$. From the first two equations, we have that

$$4x_1^3 = -\lambda = 2x_2 \qquad \Longrightarrow \qquad x_2 = 2x_1^3,$$

and from the first third equation we in a similar manner find that $x_1 = x_3$. Using these two identities in the forth equation gives that

$$x_1 + 2x_1^3 + x_1 = 1 \qquad \Longrightarrow \qquad x_1^3 + x_1 - \tfrac{1}{2} = 0.$$

This is a depressed cubic equaiton, with $p = 1$ and $q = -1/2$. It can now be verified that $\Delta := q^2/2 + p^3/27 > 0$, and using the formulas to compe the real root (using a calculator) gives $x_1 \approx 0.423853799$. This means that $x_2 \approx 0.15245004$ and $\lambda \approx -0.3045848037$. Rounding to 3 significant digits, this means that

$$x^* \approx \begin{bmatrix} 0.424 \\ 0.152 \\ 0.424 \end{bmatrix}$$

and $\lambda \approx -0.305$

(To formally prove that a globally optimal solution to the problem does exist, before explicitly finding one using the KKT conditions: Note that f is continuous, that the feasible region is a closed set (why?), and that f is weakly coercive on the feasible region. To see the last point, note that f is weakly coercive on $\mathbb{R}^3$, since for any sequence $\{x_n\}$ such that $\|x_n\| \to \infty$ we have that $f(x_n) \to \infty$, and therefore it is also weakly coercive on the feasible region. By Weierstrass' theorem, the problem therefore has (at least) one globally optimal solution.)

Question 7

(Interior penalty methods)

(1p) a) It should be continuous, map from $\mathbb{R}_-$ to $\mathbb{R}_+$, and be such that $\phi(s_k) \to \infty$ as $k \to \infty$ for any negative sequence $\{s_k\}$ that converges to 0. The only function with these properties is ϕ_5.

(2p) b) To show that it is valid output from an interior penalty method, for $k = 0, 1, 2$, we verify that $x(\nu^k)$ fulfills $g(x(\nu^k)) < 0$, and it is optimal to (P_{ν^k}).

The fact that $g(x(\nu^k)) < 0$ is easily verified.

There are multiple ways to verify that $x(\nu^k)$ is optimal to (P_{ν^k}). One such way: first note that f is convex and that g is affine and hence convex. This means that (P) is a convex problem. Moreover, $\phi(g(x))$ is also convex on the set $\{x \in \mathbb{R}^2 \mid g(x) < 0\}$ (why?). This means that (P_ν) is convex (why?). The output is thus valid if the gradient of the cost function in (P_ν) is approximately equal to zero in these points (why?). To this end, the gradient is

$$\nabla f(x) + \nu\phi'(g(x))\nabla g(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \nu \frac{1}{(2x_1 - x_2 + 30)^2} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Plugging in the values of ν^k and $x(\nu^k)$ and computing the gradient (on a clacu-lator) gives

$$\begin{bmatrix} 0.0024 \\ -0.0010 \end{bmatrix}, \qquad \begin{bmatrix} -0.0025 \\ 0.0013 \end{bmatrix}, \qquad \begin{bmatrix} -0.023 \\ 0.011 \end{bmatrix},$$

which are all "reasonably close to zero".