Chalmers/GU
Mathematics sciences

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	25–08–19
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.
Examiner:	Axel Ringh (073 708 23 73)

Exam instructions

When you answer the questions

Use generally valid theory and methods.
State your methodology carefully.
Graphical solutions must be motivated with calculations.

Only write on one page of each sheet. Do not use a red pen.
Do not answer more than one question per page.

(3p) Question 1

(Unconstrained optimization)

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x) = x_1^2 - x_1 x_2 - x_2^2 + 2\sin(x_1) + \frac{1}{8}x_2^4$$

and consider the unconstrained optimization problem

$$\text{minimize } f(x).$$

Is $x^{(1)} = [0, 0]^T$ a locally optimal solution to the problem? Is $x^{(2)} = [0, 2]^T$ a locally optimal solution to the problem? Why or why not? Motivate carefully.

Question 2

(the Karush-Kuhn-Tucker conditions)

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 $i = 1, \dots, m$.

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

(1p) b) For this subquestion, assume that x^1 is a locally optimal solution to (P), and that the linear independence constraint qualifier holds in x^1. In such a case, is it guaranteed that $x^{(1)}$ satisfies the KKT conditions? Why or why not? Motivate carefully.

(1p) c) For this subquestion, assume instead that both $x^{(2)}$ and $x^{(3)}$ satisfy the KKT conditions for (P), and that these two points are the only KKT points for the problem. Moreover, assume that Abadie's constraint qualifier holds in both points, and that the problem has at least one globally optimal solution. In such a case, is it guaranteed that at least one of the two points is the globally optimal solution to (P)? Why or why not? Motivate carefully.

Question 3

(LP duality)

Consider the linear programming problem

$$\begin{aligned} \text{minimize} \quad & x_1 + 2x_2 - x_3 \\ \text{subject to} \quad & x_1 \qquad\quad - 2x_3 \leq 1 \\ & \quad - x_2 + 2x_3 \leq 0 \\ & x_1 + 3x_2 + 5x_3 = 5 \\ & x_1 \geq 0, \;\; x_3 \leq 0. \end{aligned}$$

(2p) a) State the dual problem.

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

$$x^* = \begin{bmatrix} 0 \\ \frac{5}{3} \\ 0 \end{bmatrix}.$$

Give an optimal solution to the dual problem.

Question 4

(Lagrangian duality and global optimality conditions)

Consider the problem

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

(2p) a) Perform a Lagrange relaxation of both constraints, and derive the dual function $q(\mu)$ explicitly.

Hint: For $X := \mathbb{R}^2$, an explicit solution to $\min_{x \in X} \mathscr{L}(x, \mu)$ can be found in this case as a function of μ .

(1p) b) Using the computations in a), with $X := \mathbb{R}^2$, verify that $\hat{x} = [1, 0]^T$ and $\hat{\mu} = [2, 0]^T$ satisfy the global optimality conditions.

Question 5

(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 $f : \mathbb{R}^3 \to \mathbb{R}$ be defined by

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

and let $S = \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 \leq 1\}$.

Claim: The optimization problem

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

attains a global minimum.

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

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

Claim: $p = [0, 1]^T$ is a subgradient in the point $\tilde{x} = [1, 1]^T$.

(1p) c) Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$\begin{aligned} f(x) &= x_1^4 - x_1 x_2 + x_2^2, \\ g(x) &= \cos(x_1) + x_2^2 - 9, \end{aligned}$$

respectively, and consider the problem

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0. \end{cases}$$

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

(3p) Question 6

(Exterior penalty method)

Consider the problem

$$\begin{cases} \text{minimize} & f(x), \\ \text{subject to} & h(x) = 0, \\ & x \in \mathbb{R}^2, \end{cases}$$

where

$$f(x) = -x_1^2 + x_2^2,$$
$$h(x) = x_1^3 - x_2.$$

We will consider the exterior penalty method with penalty function $\psi(s) = s^2$.

State the penalty transformed problem that needs to be solved in each iteration of the exterior penalty method. Find all stationary points of the penalty transformed problem, that is, compute the stationary points $x^*(\nu)$ as functions of ν. Finally, for the stationary points, compute $\lim_{\nu \to \infty} x^*(\nu)$. State what you know about these limits.

Hint: An equation of the form $z(z^4 - a) = 0$, where $a > 0$, has three real solutions: $z = 0$ and $z = \pm a^{1/4}$.

Question 7

(Separation Theorem)

The separation theorem is an important theorem in convex analysis and optimization. Somewhat informally, it states that for a closed and convex set $C \subset \mathbb{R}^n$ which is nonempty, if we take a point that is not in C, we can find a hyperplane that separates the space into two halfspaces so that the point is on one side of the hyperplane and C is on the other side of the hyperplane.

(1p) a) Give the formal statement of the separation theorem.

(2p) b) Prove the separation 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.

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 25–08–19
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

(Unconstrained optimization)

We start by computing the gradient and the Hessian of f. This gives

$$\nabla f(x) = \begin{bmatrix} 2x_1 - x_2 + 2\cos(x_1) \\ -x_1 - 2x_2 + \frac{1}{2}x_2^3 \end{bmatrix}, \quad \nabla^2 f(x) = \begin{bmatrix} 2 - 2\sin(x_1) & -1 \\ -1 & -2 + \frac{3}{2}x_2^2 \end{bmatrix}.$$

For $x^{(1)}$, we get

$$\nabla f(x^{(1)}) = \begin{bmatrix} 0 - 0 + 2 \\ -0 - 0 + 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and hence $x^{(1)}$ is not a locally optimal solution.

For $x^{(2)}$, we get

$$\nabla f(x^{(2)}) = \begin{bmatrix} 0 - 2 + 2 \\ -0 - 4 + 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and $x^{(2)}$ could therefore be a locally optimal solution. To verify if it is, we check the sufficient condition for locally optimality: that the Hessian in the point is positive definite. To this end,

$$\nabla^2 f(x^{(2)}) = \begin{bmatrix} 2 - 0 & -1 \\ -1 & -2 + \frac{3}{2}2^2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}.$$

Computing the eigenvalues of this matrix, we find that $\lambda = 3 \pm \sqrt{2} > 0$ and hence the matrix is positive definite. Therefore, $x^{(2)}$ is a locally optimal solution.

Note: The positive definiteness of the matrix can also be asserted in other ways. For example, by noting that the matrix is strictly diagonally dominant. Or by noting that the leading principles minors are all positive.

Question 2

(the Karush-Kuhn-Tucker conditions)

(1p) a) For feasible points $x \in \mathbb{R}^n$, i.e., x such that $g_i(x) \leq 0$, $i = 1, \ldots, m$, the KKT conditions are

$$\nabla f(x) + \sum_{i=1}^{m} \mu_i \nabla g_i(x) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$
$$\mu_i \geq 0, \quad i = 1, \ldots, m$$
$$\mu_i g_i(x) = 0, \quad i = 1, \ldots, m.$$

(1p) b) Since the LICQ holds, Abadie's CQ holds. If Abadie's CQ holds, then KKT conditions are necessary conditions for local optimality. Therefore, since $x^{(1)}$ is assumed to be locally optimal, $x^{(1)}$ is a KKT point.

(1p) c) No it is not guaranteed that either $x^{(2)}$ or $x^{(3)}$ is a globally optimal solution. A globally optimal solution can be in a point where Abadie's constraint qualifier does not hold, and hence does not have to be a KKT point.

Question 3

(LP duality)

(2p) a) The dual problem is

$$\begin{array}{rrrrl} \text{maximize} & y_1 & & +5y_3 & \\ \text{subject to} & y_1 & & +\ y_3 & \leq 1 \\ & & -\ y_2 & +3y_3 & = 2 \\ & -2y_1 & +2y_2 & +5y_3 & \geq -1 \\ & y_1, y_2 \leq 0. & & & \end{array}$$

(1p) b) Evaluating the first and second constraint of the primal problem in the point x^*, we see that neither of them is active. By complementarity, that means that for an optimal dual solution we must have $y_1^* = y_2^* = 0$. Solving the equality constraint in the dual problem we therefore get $y_3^* = \frac{2}{3}$. It is easily verified that this is feasible to the two remaining (inequality) constraints in the dual. Since x^* and y^* are feasible in the primal and dual problem, respectively, and since complementarity holds, we know that they are optimal.

Question 4

(Lagrangian duality and global optimality conditions)

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

$$q(\mu) = \begin{array}{rl} \min & -2x_1^2 - x_2^2 + \mu_1(x_1^2 + x_2^2 - 1) + \mu_2(-x_2) \\ \text{subject to} & x \in X \end{array}$$

$$= \begin{cases} -\infty & \text{if } \mu_1 < 2, \\ -\frac{\mu_2^2}{4} - 2 & \text{if } \mu_1 = 2 \\ -\frac{\mu_2^2}{4(\mu_1 - 1)} - \mu_1 & \text{if } \mu_1 > 2 \end{cases}$$

where the minimizers of the Lagrangian for $\mu_1 = 2$ are attained at $x = [a, \mu_2/2]^T$ for any $a \in \mathbb{R}$, and for $\mu_1 > 2$ at $x = [0, \mu_2/(2(\mu_1 - 1))]^T$.

(1p) b) Let $X := \mathbb{R}^2$, $g_1(x) = x_1^2 + x_2^2 - 1$, and $g_2(x) = -x_2$. A point $(x^*, \mu^*) \in \mathbb{R}^2 \times \mathbb{R}^2$ is said to satisfy the global optimality conditions if

$$\begin{aligned} x^* \in \arg\min_X \ \mathcal{L}(x, \mu^*), \\ x^* \in X, \ g_1(x^*) \leq 0, \ \text{and } g_2(x^*) \leq 0, \\ \mu^* \geq 0, \\ \mu_1^* g_1(x^*) = 0, \ \text{and } \mu_2^* g_2(x^*) = 0. \end{aligned}$$

We were told to consider the point $\hat{x} = [1, 0]^T$ and $\hat{\mu} = [2, 0]^T$. From a), we know that $\arg\min_X \ \mathcal{L}(x, \hat{\mu}) = \{x = [a, \hat{\mu}_2/2]^T \mid a \in \mathbb{R}\}$ and thus $\hat{x} \in \arg\min_X \ \mathcal{L}(x, \hat{\mu})$. Moreover, $g_1(\hat{x}) = 1^2 + 0^2 - 1 = 0 \leq 0$ and $g_2(\hat{x}) = -0 = 0 \leq 0$. Next, $\hat{\mu}_1 = 1 \geq 0$ and $\hat{\mu}_2 = 0 \geq 0$. Finally, $\hat{\mu}_1 g_1(\hat{x}) = 2 \cdot 0 = 0$ and $\hat{\mu}_2 g_2(\hat{x}) = 0 \cdot 0 = 0$. This shows that $\hat{x} = [1, 0]^T$ and $\hat{\mu} = [2, 0]^T$ satisfy the global optimality conditions.

Question 5

(True or False)

(1p) a) True. f is continuous and S is nonempty, closed, and bounded. So by Weierstrass' theorem the problem attains a globally optimal solution.

(1p) b) False. Since f is continuously differentiable in $\tilde{x}$, $\partial f(\tilde{x}) = \{\nabla f(\tilde{x})\}$. But

$$\nabla f(x) = \begin{bmatrix} 2x_1 - x_2 \\ -x_1 + 2x_2 \end{bmatrix} \qquad \Longrightarrow \qquad \nabla f(\tilde{x}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 1 \end{bmatrix} = p.$$

(1p) c) False. Newton's method, as it is taught in the course, can only be used to solve unconstrained problems.

(3p) Question 6

(Exterior penalty method)

The penalty transformed problem that needs to be solved in each iteration of the exterior penalty method is

$$\min_{x\in\mathbb{R}^2} f(x)+\nu\psi(h(x)) = \min_{x\in\mathbb{R}^2} -x_1^2+x_2^2+\nu(x_1^3-x_2)^2 = \min_{x\in\mathbb{R}^2} -x_1^2+x_2^2+\nu x_1^6-2\nu x_1^3x_2+\nu x_2^2.$$

Let $F_\nu(x) := -x_1^2 + x_2^2 + \nu x_1^6 - 2\nu x_1^3 x_2 + \nu x_2^2$ denote the cost function of the penalty transformed problem. Stationary points to the penalty transformed problem are points such that $\nabla F_\nu(x) = 0$.

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \nabla F_\nu(x) = \begin{bmatrix} -2x_1 + 6\nu x_1^5 - 6\nu x_1^2 x_2 \\ 2x_2 - 2\nu x_1^3 + 2\nu x_2 \end{bmatrix}.$$

The second equation gives that $x_2 = \frac{\nu}{1+\nu}x_1^3$, and plugging this into the first equation gives that

$$0 = -2x_1+6\nu x_1^5-6\nu x_1^2\frac{\nu}{1+\nu}x_1^3 = 2x_1\left(-1+3\nu x_1^4-\frac{3\nu^2}{1+\nu}x_1^4\right) = 2x_1\left(\frac{3\nu}{1+\nu}x_1^4-1\right).$$

The (real) solutions to this equation are $x_1(\nu) = 0$, and $x_1(\nu) = \pm(\frac{1+\nu}{3\nu})^{1/4}$ which means that the stationary points are

$$x^*(\nu) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad x^*(\nu) = \pm \begin{bmatrix} \left(\frac{1}{3}\right)^{1/4}\left(\frac{1+\nu}{\nu}\right)^{1/4} \\ \left(\frac{1}{3}\right)^{3/4}\left(\frac{\nu}{1+\nu}\right)^{1/4} \end{bmatrix}.$$

The limit, as $\nu \to \infty$, of these points are

$$\lim_{\nu\to\infty} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \lim_{\nu\to\infty} \pm \begin{bmatrix} \left(\frac{1}{3}\right)^{1/4}\left(\frac{1+\nu}{\nu}\right)^{1/4} \\ \left(\frac{1}{3}\right)^{3/4}\left(\frac{\nu}{1+\nu}\right)^{1/4} \end{bmatrix} = \pm \begin{bmatrix} \left(\frac{1}{3}\right)^{1/4} \\ \left(\frac{1}{3}\right)^{3/4} \end{bmatrix}.$$

Since all functions are C^1, since $\psi'(s) = 2s \geq 0$ for all $s \geq 0$, since all of the above limit points are feasible to the original problem, and since the LICQ holds in all of the above limit points, we know that they are all KKT points to the original problem (see Theorem 13.4 in course book).

Question 7

(Separation Theorem)

(1p) a) See Theorem 4.29 in the book.

(2p) b) See the proof of Theorem 4.29 in the book.
