Chalmers/GU
Mathematics sciences

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	23–08–15
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.

(3p) Question 1

(The Frank-Wolf algorithm)

Consider the problem

$$\begin{aligned} \text{minimize} \quad & x_1^2 x_2^3 + x_2^2 \\ \text{subject to} \quad & -x_1 + x_2 \leq 3 \\ & 2x_1 + x_2 \leq 12 \\ & x_1 - 2x_2 \leq 2 \\ & x_1 \geq 0 \\ & x_2 \geq 0. \end{aligned}$$

Start in the point $x_1 = 1$ and $x_2 = 1$ and perform one iteration in the Frank-Wolfe algorithm. More precisely, this means that you have to: write down the correct subproblem for computing the search direction; solve the subproblem using a generally valid solution method for this type of problem; compute the step length using exact line search; perform the update and write down the new point. Carefully motivate all steps and conclusions.

Question 2

(LP duality)

Consider the linear programming problem

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \geq b, \\ & x \leq u, \end{aligned}$$

where a $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $u \in \mathbb{R}^n$.

(2p) a) Construct the LP dual of this primal linear programming problem.

(1p) b) Construct the dual LP of the dual LP.

Question 3

(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 $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable.

Claim: If $\nabla f(x_0)^T p \geq 0$, then p cannot be a descent direction with respect to f at the point x_0.

(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

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

respectively.

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

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

(1p) c) Let $f_1 : \mathbb{R}^n \to \mathbb{R}$ and $f_2 : \mathbb{R}^n \to \mathbb{R}$ be two convex functions.

Claim: The function $f(x) := f_1(x) + f_2(x)$ is convex.

Question 4

(Weierstrass' Theorem)

For each of the following functions f_i, $i = 1, 2, 3$, motivate carefully whether or not a global minimum is attained on the corresponding set S_i, $i = 1, 2, 3$.

(1p) a) $f_1 : \mathbb{R}^n \to \mathbb{R}$ given by $f_1(x) = \sum_{\ell=1}^{n} x_\ell^{2\ell+1}$ and $S_1 = \{x \in \mathbb{R}^n \mid 0 \leq x_\ell \leq 1,\ \ell = 1, \dots, n\}$.

(1p) b) $f_2 : \mathbb{R} \to \mathbb{R}$ given by

$$f_2(x) = \begin{cases} 1 & \text{if } x = 0 \\ |x| & \text{if } |x| > 0 \end{cases}$$

and $S_2 = \{x \in \mathbb{R} \mid -10 \leq x \leq 10\}$.

(1p) c) $f_3 : \mathbb{R}^2 \to \mathbb{R}$ given by $f_3(x) = x_1^2 + e^{\cos(2\pi x_2)}$ and $S_3 = \{x \in \mathbb{R}^2 \mid 0 \leq x_2 \leq 100\}$.

(3p)
Question 5

(Optimality conditions for unconstrained problems and Convex quadratic programming)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = \frac{1}{2}x^T Q x + c^T x$, where $c \in \mathbb{R}^n$ and the symmetric matrix $Q \in \mathbb{R}^{n \times n}$ is positive semi-definite. Consider the problem

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

Under what conditions on Q and c does this problem have a globally optimal solution?

Question 6

(Global optimality conditions)

Let $X \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, and $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \dots, m$. Consider the problem

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

Let f^* be the infimum for problem (P), and assume that $-\infty < f^* < \infty$. The Lagrangian to (P) is defined as

$$\mathscr{L}(x, \mu) = f(x) + \sum_{i=1}^{m} \mu_i g_i(x),$$

and a vector $\mu^* \in \mathbb{R}^m$ is called a Lagrangian multiplier if $\mu^* \geq 0$ and

$$f^* = \begin{array}{ll} \text{infimum} & \mathscr{L}(x, \mu^*) \\ \text{subject to} & x \in X. \end{array}$$

(1p) a) State the global optimality conditions for problem (P).

(2p) b) Prove the following theorem:

THEOREM: (x^*, μ^*) *satisfies the global optimality conditions if and only if* x^* *is globally optimal to* (P) *and* μ^* *is a Lagrangian multiplier.*

(3p) Question 7

(Modelling)

A company is planning for where to locate a number of factories to produce a certain product to sell. The goal of the planning is to find an arrangement of factories that makes the total production as cheap as possible.

More specifically, the company wants to construct two types of factories. After an initial investigation, they have found that the first type of factory can be placed in any of the locations a_i for $i \in I := \{1, \dots, n_i\}$, and to construct such a factory in location a_i costs $e_i > 0$ SEK. The second type of factory can be placed in any of the locations b_j, for $j \in J := \{1, \dots, n_j\}$, and to construct such a factory in b_j costs $f_j > 0$ SEK.

The first type of factory produces a certain raw material. This raw material is needed in the second type of factory. The second type of factory produces the end product that the company sells. The costs for production of both the raw material and the end product are not taken into account in the model.

Nevertheless, the cost for transporting raw material from factories of type one to factories of type two must be taken into account in the model. The cost for transporting raw material from a factory of type one in location a_i to a factory of type two in location b_j is $c_{ij} > 0$ SEK/kg of raw material transported, for $i \in I$ and $j \in J$. However, transportation between two factories can only be done if a road between the two factories is constructed. Constructing a road between a factory of type one located in a_i and a factory of type two located in b_j costs $p_{ij} > 0$ SEK, for $i \in I$ and $j \in J$.

A factory of the first type in location a_i produces $h_i > 0$ kg of the raw material, which can be transported to factories of the second type. Not all of the raw material produced in a factory of type one must to be transported to a factory of type two. Furthermore, if a factory of the second type is constructed in location b_j, it must receive exactly $k_j > 0$ kg of raw material. From this raw material, a factory of the second type in location b_j produces $\ell_j > 0$ units of the end product. The company wants to produce at least $L > 0$ units of the end product.

Help the company by formulating an integer linear program, using the information given above, that minimizes the cost for constructing factories, constructing roads, and sending the raw material from factories of type one to factories of type two.

Hint 1: Consider introducing three sets of binary variables that models: where factories of type one are constructed; where factories of type two are constructed; which roads between factories of type one and type two that are constructed.

Hint 2: Consider introducing a set of variables that determines the amount of flow of raw material from factories of type one to factories of type two. Make sure to formulate the model so that there can be no flow between two factories if a road between them has not been constructed.

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621 NONLINEAR OPTIMISATION

Date: 23–08–15
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

(The Frank-Wolf algorithm)

Subproblem for computing search direction:
Let $f(x) = x_1^2 x_2^3 + x_2^2$, which gives

$$\nabla f(x) = \begin{bmatrix} 2x_1 x_2^3 \\ 3x_1^2 x_2^2 + 2x_2 \end{bmatrix} \quad \Longrightarrow \quad \nabla f\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 5 \end{bmatrix}.$$

The search direction p is given by $p = y - [1,\ 1]^T$, where y is the optimal solution to the subproblem

$$\begin{aligned} \text{minimize} \quad & 2y_1 + 5y_2 \\ \text{subject to} \quad & -y_1 + y_2 \leq 3 \\ & 2y_1 + y_2 \leq 12 \\ & y_1 - 2y_2 \leq 2 \\ & y_1, y_2 \geq 0. \end{aligned}$$

Transforming the latter to standard form gives

$$\begin{aligned} \text{minimize} \quad & 2y_1 + 5y_2 \\ \text{subject to} \quad & -y_1 + y_2 + s_1 = 3 \\ & 2y_1 + y_2 + s_2 = 12 \\ & y_1 - 2y_2 + s_3 = 2 \\ & y_1, y_2, s_1, s_2, s_3 \geq 0. \end{aligned}$$

To solve the linear program with Simplex, we see that it is possible to start with s_1, s_2, and s_3 as basic variables (motivate!).

Simplex - Iteration 1:
With $x_B = [s_1,\ s_2,\ s_3]^T$ and $x_N = [x_1,\ x_2,]^T$,

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} -1 & 1 \\ 2 & 1 \\ 1 & -2 \end{bmatrix}, \quad c_B^T = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \quad c_N^T = \begin{bmatrix} 2 & 5 \end{bmatrix}.$$

$x_B = B^{-1}b = [3,\ 12,\ 2]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [2,\ 5] \geq 0$. Hence the point $y^* = [0,\ 0]^T$ is optimal to the subproblem.

Line search and update:
The search direction is $p = [0,\ 0]^T - [1,\ 1]^T = [-1,\ -1]^T$, and the exact line search problem becomes

$$\begin{aligned} \text{minimize} \quad & f\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ -1 \end{bmatrix}\right) = (1-\alpha)^5 + (1-\alpha)^2 \\ \text{subject to} \quad & \alpha \in [0, 1]. \end{aligned}$$

To optimal solution to this problem is $\alpha^* = 1$ (motivate!), and hence the new point is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Question 2

(LP duality)

(2p) a) The LP dual is given by

$$\begin{aligned} \text{maximize} \quad & b^T y - u^T z \\ \text{subject to} \quad & A^T y - z = c, \\ & y \geq 0, \\ & z \geq 0, \end{aligned}$$

where $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$.

(1p) b) The dual of the dual is the primal. So it is

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \geq b, \\ & x \leq u. \end{aligned}$$

Question 3

(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) False. A counterexample is given by $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, $x_0 = 0$, and $p = -1$. $\nabla f(x_0)^T p = 0 \geq 0$ but $f(x_0 + \alpha p) = -\alpha^3 < 0$ for all $\alpha > 0$.

(1p) b) True. The relaxed problem obtained by removing constraint g_1 is convex (motivate why!), and $x^* = [1, 1]^T$ is the unique globally optimal solution to the relaxed problem (need to show that!). Since x^* is feasible to both the original and the relaxed problem, and since the objective function is the same in both problems, x^* is globally optimal to the original problem (motivate!).

(1p) c) True. The sum of convex functions is convex. For all $x, y \in \mathbb{R}^n$ and all $\alpha \in [0, 1]$

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &= f_1(\alpha x + (1-\alpha)y) + f_2(\alpha x + (1-\alpha)y) \\ &\leq \alpha f_1(x) + (1-\alpha) f_1(y) + \alpha f_2(x) + (1-\alpha) f_2(y) \\ &= \alpha f(x) + (1-\alpha) f(y), \end{aligned}$$

where the inequality follows by convexity of f_1 and f_2.

Question 4

(Weierstrass' Theorem)

For each of the following functions f_i, $i = 1, 2, 3$, motivate carefully whether or not a global minimum is attained on the corresponding set S_i, $i = 1, 2, 3$.

(1p) a) f_1 is continuous, and S_1 is compact. Therefore, by Weierstrass' Theorem there exists a set of globally optimal solutions.

(1p) b) The infimum of f_2 over S_2 is 0 (why? motivate!), but the value is not attained for any x. (Note: f_2 is not lower semi-continuous.)

(1p) c) f_3 is continuous, but S_3 is not compact. However, every direction of unboundedness of S is such that $x_1 \to \pm\infty$ and hence f_1 is weakly coercive with respect to S_1 (why? motivate!). Therefore, by Weierstrass' Theorem there exists a set of globally optimal solutions.

Question 5

(Optimality conditions for unconstrained problems and Convex quadratic programming)

The gradient and Hessian of f are given by

$$\nabla f(x) = Qx + c, \qquad \nabla^2 f(x) = Q.$$

Since Q is positive semi-definite, f is a convex function (motivate! Theorem 3.49). Since f is convex, a point x is globally optimal if and only if $\nabla f(x) = 0$ (motivate! Theorem 4.18). Therefore, the problem has a globally optimal solution if and only if the system of linear equations

$$Qx = -c$$

is solvable, i.e., if c lies in the column space of Q.

Note: if Q is positive definite, it is invertible and thus $Qx = -c$ has a unique solution for all c (if Q is positive definite, then f is also strictly convex). Since Q is positive semi-defintie, if it is not positive definite then it has at least one eigenvalue that is zero. In this case, Q has a non-trivial kernel and $Qx = -c$ has a solution if and only if c lies in the column space of Q. If c does not lay in the column space of Q, then there exists a vector v such that $Qv = 0$ and $c^T v < 0$. In this case, for $k \in \{1, 2, 3, \ldots\}$ and $x^{(k)} = kv$ we have $f(x^{(k)}) = (x^{(k)})^T Q x^{(k)} + c^T x^{(k)} = k(c^T v)$, and hence the sequence $\{x^{(k)}\}_{k=1}^{\infty}$ shows that the problem is unbounded from below, i.e., $\lim_{k\to\infty} f(x^{(k)}) = -\infty$. Therefore, the problem admits no optimal solution in this case.

Question 6

(theory question - global optimality conditions)

(1p) a) See Theorem 6.8 in the book.

(2p) b) See Theorem 6.8 in the book.

(3p) Question 7

(modelling)

For each $i \in I$, introduce

$$x_i = \begin{cases} 1 & \text{factory of type one constructed in } a_i \\ 0 & \text{else,} \end{cases}$$

and each $j \in J$, introduce

$$y_j = \begin{cases} 1 & \text{factory of type two constructed in } b_j \\ 0 & \text{else.} \end{cases}$$

Moreover, for all $i \in I$ and all $j \in J$, also introduce

$$z_{ij} = \begin{cases} 1 & \text{if link between factory of type one in } a_i \text{ and factory of type two in } b_j \text{ is constructed,} \\ 0 & \text{else,} \end{cases}$$

and v_{ij} = weight (in kg) of raw material sent from factory of type one in a_i to factory of type two in b_j. Finally, let M be a large positive constant ($M \geq \max\{\max_{i \in I} h_i, \max_{j \in J} k_j\}$ suffices). An integer linear program can be formulated as

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{I} e_i x_i + \sum_{j=1}^{J} f_j y_j + \sum_{i=1}^{I}\sum_{j=1}^{J} c_{ij} v_{ij} + \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} z_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{J} v_{ij} \leq h_i x_i, \ i \in I, && \text{(constructed type one can send maximum amount)} \\
& \sum_{i=1}^{I} v_{ij} = k_j y_j, \ j \in J, && \text{(constructed type two wants exact amount)} \\
& v_{ij} \leq M z_{ij}, \quad i \in I, \ j \in J, && \text{(can only send on constructed roads)} \\
& \sum_{j=1}^{J} \ell_j y_j \geq L, \ j \in J, && \text{(constructed type two must produce sufficient total)} \\
& x_i \in \{0,1\}, \quad i \in I, \\
& y_j \in \{0,1\}, \quad j \in J, \\
& z_{ij} \in \{0,1\}, \quad i \in I, \ j \in J, \\
& v_{ij} \geq 0, \quad i \in I, \ j \in J.
\end{aligned}$$

Note that we do note need constraints to enforce that $z_{ij} = 0$ if $x_i = 0$ or $y_j = 0$, since the goal is to minimize the cost and $p_{ij} > 0$.