TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	23–01–04
Time:	8^{30}–13^{30}
Aids:	Text memory-less 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 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 simplex method)

Consider the following linear program:

$$
\begin{array}{rlr}
\text{minimize} & 8x_1 - x_2 - 2x_3 & \\
\text{subject to} & 3x_1 + 2x_2 & \leq 21, \\
& -3x_1 + x_2 + x_3 \leq 7, \\
& x_1,\ x_2,\ x_3 & \geq 0.
\end{array}
$$

Solve it using the simplex method (phase II), and start with x_1 and x_3 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 2

(Lagrangian duality)

Consider the problem

$$
\begin{array}{rl}
\text{minimize} & \displaystyle\sum_{i=1}^{n} a_i \Big(x_i \log(x_i) - x_i \Big), \\
\text{subject to} & \displaystyle\sum_{i=1}^{n} b_i x_i = 1, \\
& x_i \geq 0, \quad i = 1, \ldots, n,
\end{array}
$$

where log denotes the natural logarithm, $a_i > 0$ for all $i = 1, \ldots, n$, $b_i > 0$ for all $i = 1, \ldots, n$, and $0 \log(0)$ is defined to be equal to 0. Lagrangian relax the constraint $\sum_{i=1}^{n} b_i x_1 = 1$ and derive the Lagrangian dual problem. You have to wright the dual problem explicitly, but you do not have to solve it.

Note: The fact that $0 \log(0)$ is defined to be equal to 0 makes the cost function continuous on the entire domain $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1, \ldots, n\}$.

Question 3

(global convergence of exterior penalty method)

Consider the problem

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

and let

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

Also consider the transformed problem

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

where

$$
\check{\chi}_S(x) = \sum_{i=1}^{m} \psi(\max\{0, g_i(x)\}) + \sum_{j=1}^{\ell} \psi(h_j(x)).
$$

(1p)   a) Define $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \ge 0\}$. What condition must the function $\psi : \mathbb{R} \to \mathbb{R}_+$ satisfy for us to call (P$_\nu$) an exterior penalty transformation of (P)?

(2p)   b) A function $\psi : \mathbb{R} \to \mathbb{R}_+$ that fulfills the conditions asked for in part a) is called an exterior penalty function. Prove the following theorem.

THEOREM: *Let ψ be an exterior penalty function, and assume that* (P) *has at least one globally optimal solution. For each value of ν, let x_ν^* be a globally optimal solution to* (P$_\nu$). *Then every limit point of the sequence $\{x_\nu^*\}$, $\nu \to \infty$, is a globally optimal solution to* (P).

Hint: The following result might be useful. You may use it without proving it.

LEMMA: *Let $x_{\nu_1}^*$ and $x_{\nu_2}^*$ be globally optimal to* (P$_\nu$) *for penalty parameters ν_1 and ν_2, respectively. If $\nu_1 \le \nu_2$, then $f(x_{\nu_1}^*) \le f(x_{\nu_2}^*)$.*

Question 4

(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 twice contiunously differentiable.

Claim: For the unconstrained optimization problem $\min_{x \in \mathbb{R}^n} f(x)$, the conditions i) $\nabla f(x^*) = 0$, and ii) $\nabla^2 f(x^*)$ is positive semi-definite, are sufficient for x^* to be a local minimum.

(1p)   b) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable.

Claim: For p to be a descent direction with respect to f at the point x_0, it is sufficient that $\nabla f(x_0)^T p < 0$.

(1p)   c) Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable and strictly convex.

Claim: The problem to minimize f over $\mathbb{R}^n$ has a unique optimal solution.

Question 5

(the Karush-Kuhn-Tucker (KKT) conditions)

Consider the problem

$$
\begin{cases}
\text{minimize} & f(x), \\
\text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, 3, \\
& x \in \mathbb{R}^2,
\end{cases}
$$

where

$$
\begin{aligned}
f(x) &= -(x_1 - 1)^2, \\
g_1(x) &= -x_1^3 + x_2, \\
g_2(x) &= x_1^2 + x_2^2 - 2, \\
g_3(x) &= -x_2 \leq 0.
\end{aligned}
$$

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

(1p)   b) Find all KKT points, i.e., all points x that satisfy the KKT conditions. Solutions based on graphical considerations are allowed, but they need to be supplemented with exact mathematical expressions and calculations motivating the conclusions.

Hint: The only point for which g_1 and g_2 can both be active is $x_1 = x_2 = 1$.

(1p)   c) Which of the KKT points have smallest objective function value? Is this KKT point globally optimal?

(3p) **Question 6**

(modelling)

A friend of yours, we can call the person AR, is planning a road trip. AR has decided to visit a number of cities $1, \ldots, n$, and the trip is such that from city i AR will drive to city $i + 1$, for $i = 1, \ldots, n - 1$.

For the trip, AR has rented an electric car. The car is picked up in city 1 and returned in city n, and AR is now planning how to charge the car during the trip in order to minimize the cost of charging. More specifically, the car battery has a maximum energy capacity of K kWh (kilowatt-hours), and when picking up the car the battery is fully charged. When returning the car, the battery needs to be at least 60% charged. Moreover, to drive from city i to city $i + 1$ the total energy needed for the car is e_i kWh, for $i = 1, \ldots, n - 1$. In each city, AR will be able to charge the car between 0 and a maximum of t_i hours, for $i = 2, \ldots, n$. The power output from the charging stations in each city is p_i kW (kilowatts), and the cost for charging is c_i SEK/kWh, for $i = 2, \ldots, n$. In particular, note that AR can charge the car in the last city n before returning it to the car rental shop.

Help AR finding the cheapest way to charge the car, while also being able to complete the whole trip. Formulate it as a linear programming problem.

Hint 1: Consider introducing two sets of variables. One set that represents the amount of energy charged in a city, and one that represents the amount of energy in the battery at different points/times.

Hint 2: Charging a battery at a charging station with power output p kW for t hours results in a total energy in the battery of pt kWh.

(3p) **Question 7**

(unconstrained optimization - Newton's method with Levenberg-Marquardt modification)

Consider the unconstrained optimization problem

$$\text{minimize} \quad f(x)$$
$$\text{subject to} \quad x \in \mathbb{R}^2,$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is given by

$$f(x) = \frac{1}{4}x_1^4 - 3x_1^2x_2^2 + 2x_1x_2^3 + x_2^4 - x_1 + x_2^2.$$

Starting in the point $[x_1, \ x_2]^T = [0, \ 0]^T$, perform one iteration in Newton's method using the Levenberg-Marquardt modification. In particular, select the modification parameter $\gamma > 0$ as the smallest integer so that the conditions needed are fulfilled. Moreover, use step length $\alpha_1 = 2$.

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 23–01–04

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 simplex method)

Transforming the problem to standard form gives

$$
\begin{aligned}
\text{minimize} \quad & 8x_1 - x_2 - 2x_3 \\
\text{subject to} \quad & 3x_1 + 2x_2 + s_1 \qquad\quad = 21, \\
& -3x_1 + x_2 + x_3 + s_2 = 7, \\
& x_1,\ x_2,\ x_3,\ s_1,\ s_2 \qquad \geq 0.
\end{aligned}
$$

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

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

$x_B = B^{-1}b = [7,\ 28]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [-1/3,\ -2/3,\ 2]$, and hence $(x_N)_2 = s_1$ enters the basis. $B^{-1}N_2 = [1/3,\ 1]^T$, and for the minimum ration test we thus get $i = \arg\min_k \{\frac{7}{1/3}, \frac{28}{1}\} = \arg\min_k\{21, 28\} = 1$. This means that $(x_B)_1 = x_1$ leaves the basis.

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

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

$x_B = B^{-1}b = [7,\ 21]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [2, 1, 2]$. Hence the point $x^* = [0,\ 0,\ 7]^T$ is optimal.

(3p) Question 2

(Lagrangian duality)

Relaxing the constraint with a multiplier $\mu \in \mathbb{R}$, we get the Lagrangian

$$\mathscr{L}(x, \mu) = \sum_{i=1}^{n} a_i \Big(x_i \log(x_i) - x_i \Big) + \mu \left(\sum_{i=1}^{n} b_i x_i - 1 \right).$$

Introducing $X := \{x \in \mathbb{R}^n \mid x_i \geq 0, \ i = 1, \ldots, n\}$, the dual function is given by $q(\mu) = \inf_{x \in X} \mathscr{L}(x, \mu)$. For each fixed μ, $\mathscr{L}(x, \mu)$ is convex in x (motivate!) and X is a convex set. This means that $\inf_{x \in X} \mathscr{L}(x, \mu)$ is a convex optimization problem. If there is a point $x^* \geq 0$ such that $\nabla_x \mathscr{L}(x^*, \mu) = 0$, then x^* solves $\inf_{x \in X} \mathscr{L}(x, \mu)$ (why? motivate!). To this end, we consider the equation

$$0 = \frac{\partial \mathscr{L}(x, \mu)}{\partial x_i} = a_i \log(x_i) + \mu b_i,$$

which has a solution $x_i^* = e^{-\mu b_i / a_i} > 0$. This holds for $i = 1, \ldots, n$, and with $x^* = [x_i^*]_{i=1}^n$ we therefore have that

$$
\begin{aligned}
q(\mu) = \inf_{x \in X} \mathscr{L}(x, \mu) &= \mathscr{L}(x^*, \mu) \\
&= \sum_{i=1}^{n} a_i \Big(e^{-\mu b_i / a_i} (-\mu b_i / a_i) - e^{-\mu b_i / a_i} \Big) + \mu \left(\sum_{i=1}^{n} b_i e^{-\mu b_i / a_i} - 1 \right) \\
&= -\mu - \sum_{i=1}^{n} a_i e^{-\mu b_i / a_i}.
\end{aligned}
$$

The dual problem is thus

$$
\begin{aligned}
\text{maximize} \quad &- \mu - \sum_{i=1}^{n} a_i e^{-\mu b_i / a_i}, \\
\text{subject to} \quad &\mu \in \mathbb{R}.
\end{aligned}
$$

Question 3

(theory question - global convergence of exterior penalty method)

(1p)   a) ψ must be continuous, and $\psi(s) = 0$ if and only if $s = 0$; see Section 13.1.1 in the book.

(2p)   b) See Theorem 13.3 in the book.

Question 4

(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. The conditions are necessary for x^* to be a local minimum but not sufficient. For example, let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^3$. In $x^* = 0$ we have that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) = 0$. But x^* is not a local minimum.

(1p)   b) True. Since f is continuously differentiable, in a neighbourhood of the given point x_0 we have the Taylor series expansion

$$f(x_0 + \alpha p) = f(x_0) + \alpha \nabla f(x_0)^T p + o(\alpha).$$

Since $\nabla f(x_0)^T p < 0$ and $\lim_{\alpha \searrow 0} o(\alpha)/\alpha = 0$, there exists a sufficiently small $\delta > 0$ so that for all $\alpha \in (0, \delta]$ we have that $f(x_0 + \alpha p) < f(x_0)$, i.e., p is a descent direction with respect to f at the point x_0.

(1p)   c) False. A counterexample is given by $f(x) = e^x$.

Question 5

(the Karush-Kuhn-Tucker (KKT) conditions)

(1p)   a) The KKT conditions are

$$\nabla f(x) + \sum_{i=1}^{3} \mu_i \nabla g_i(x) = \begin{bmatrix} -2x_1 + 2 \\ 0 \end{bmatrix} + \mu_1 \begin{bmatrix} -3x_1^2 \\ 1 \end{bmatrix} + \mu_2 \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \mu_3 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mu_i \geq 0, \ i = 1, 2, 3$$
$$g_i(x) \leq 0, \ i = 1, 2, 3$$
$$\mu_i g_i(x) = 0, \ i = 1, 2, 3.$$

(1p)   b) The KKT points can be found by going over all possible combinations of constraints that can be active together:

- with no constraint active, the KKT points are given by $\nabla f(x) = 0$, which gives $x_1 = 1$ and $x_2 \in [0,1]$. So $x_1 = 1$ and $x_2 \in [0,1]$, with $\mu_1 = \mu_2 = \mu_3 = 0$ are all KKT points.

- with g_1 active, we get that we must have $\mu_1 = 0$. This in turn implies that we must have $x_1 = 1$, and since g_1 is active thus that we must have $x_2 = 1$. This (again) gives the valid KKT point $x_1 = x_2 = 1$ with $\mu_1 = \mu_2 = \mu_3 = 0$.

- with g_1 and g_2 active, the only possible point is $x_1 = x_2 = 1$. Using this, we find that we must have $\mu_1 = \mu_2 = 0$. This (again) gives the valid KKT point $x_1 = x_2 = 1$ with $\mu_1 = \mu_2 = \mu_3 = 0$.

- with g_2 active, either we must have $\mu_2 = 0$, in which case $x_1 = 1$, which in turn implies that $x_2 = 1$ (in order to be active on g_2). Or $x_2 = 0$, in which case we must have $x_1 = \sqrt{2}$ (in order to be active on g_2), and hence $\mu_2 = 1 - 1/\sqrt{2} > 0$. This (again) gives the valid KKT point $x_1 = x_2 = 1$ with $\mu_1 = \mu_2 = \mu_3 = 0$, as well as the valid KKT point $x_1 = \sqrt{2}$, $x_2 = 0$, with $\mu_2 = 1 - 1/\sqrt{2}$ and $\mu_1 = \mu_3 = 0$.

- with g_2 and g_3 active, $x_2 = 0$ and $x_1 = \sqrt{2}$ is the only possible point. This gives $\mu_2 = 1 - 1/\sqrt{2} > 0$ and $\mu_3 = 0$. Hence, it (again) gives the valid KKT point $x_1 = \sqrt{2}$, $x_2 = 0$, with $\mu_2 = 1 - 1/\sqrt{2}$ and $\mu_1 = \mu_3 = 0$.

- with g_3 active, we have $x_2 = 0$. We also find that we must have $\mu_3 = 0$ and hence that $x_1 = 1$. This (again) gives the valid KKT point $x_1 = 1$ and $x_2 = 0$, with $\mu_1 = \mu_2 = \mu_3 = 0$.

- with g_1 and g_3 active, the only feasible point is $x_1 = x_2 = 0$. However, this is not a KKT point since

$$\begin{bmatrix} 2 \\ 0 \end{bmatrix} + \mu_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \mu_3 \begin{bmatrix} 0 \\ -1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

  for any μ_1, μ_3.

(1p)   c) The smallest value among the KKT points is obtained at the point $\hat{x} = [\sqrt{2}, \ 0]^T$. But $f(\hat{x}) = -3 + 2\sqrt{2} > -3 + 2 = -1 = f(\tilde{x})$, where $\tilde{x} = [0, \ 0]^T$, so $\hat{x}$ is not globally optimal.

(3p) Question 6

(modelling)

Introduce the variables

$$x_i = \text{amount of energy (kWh) charged in city } i$$

for $i = 2, \ldots, n$, and

$$y_i = \text{amount of energy (kWh) in the battery when leaving city } i$$

for $i = 1, \ldots, n$, and where y_n is interpreted as the amount of energy (kWh) in the battery when returning the car to the rental agency. A linear program for minimizing cost of charging can be formulated as

$$\min_{\substack{x_i, \ i=2,\ldots,n \\ y_i, \ i=1,\ldots,n}} \quad \sum_{i=2}^{n} c_i x_i$$

$$
\begin{aligned}
\text{subject to} \quad & y_1 = K, && \text{(fully charged when picking up)} \\
& y_n \geq 0.6K, && \text{(minimum energy when returning)} \\
& y_i = y_{i-1} - e_{i-1} + x_i, \quad i = 2, \ldots, n, && \text{(change in energy after trip + recharge)} \\
& y_i \geq e_i, \quad i = 1, \ldots, n-1, && \text{(sufficient energy for each trip)} \\
& y_i \leq K, \quad i = 2, \ldots, n, && \text{(maximum energy in battery)} \\
& x_i \leq p_i t_i, \quad i = 2, \ldots, n, && \text{(maximum energy possible to charge)} \\
& y_i \geq 0, \quad i = 1, \ldots, n, && \\
& x_i \geq 0, \quad i = 2, \ldots, n. &&
\end{aligned}
$$

(3p) Question 7

(Unconstrained optimization - Newton's method with Levenberg-Marquardt modification)

Computing the gradient and Hessian, we have that

$$\nabla f(x) = \begin{bmatrix} x_1^3 - 6x_1 x_2^2 + 2x_2^3 - 1 \\ -6x_1^2 x_2 + 6x_1 x_2^2 + 4x_2^3 + 2x_2 \end{bmatrix}, \quad \nabla^2 f(x) = \begin{bmatrix} 3x_1^2 - 6x_2^2 & -12x_1 x_2 + 6x_2^2 \\ -12x_1 x_2 + 6x_2^2 & -6x_1^2 + 12x_1 x_2 + 12x_2^2 + 2 \end{bmatrix}.$$

This means that

$$\nabla f(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \nabla^2 f(0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

The smallest modification parameter $\gamma > 0$ so that $\nabla^2 f(0) + \gamma I$ is positive definite, and so that γ is also an integer, is thus $\gamma = 1$, which gives

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

The search direction is thus $p_1 = -(\nabla^2 f(0) + 1I)^{-1} \nabla f(0) = [1, \ 0]^T$, and the new point is

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$