Chalmers/GU
Mathematics sciences

EXAM

TMA947/MMG621 NONLINEAR OPTIMISATION

Date: 24–01–04
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 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

(Linear programming)

Consider the linear programming problem

$$\begin{aligned} \text{minimize} \quad & -4x_1 - 2x_2 - 6x_3 \\ \text{subject to} \quad & 2x_1 - x_2 + 4x_3 \leq 30 \\ & x_1 + 2x_2 + 4x_3 = 80 \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

Solve the problem using the Simplex method (phase-II), starting with x_2 and x_3 as basic variables.

Question 2

(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 $f : \mathbb{R}^2 \to \mathbb{R}$, $g_1 : \mathbb{R}^2 \to \mathbb{R}$, and $g_2 : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$\begin{aligned} f(x) &= -x_1^4 + x_1^2 x_2^2 + 0.5x_2^3 - x_1 + 3, \\ g_1(x) &= -(x_1 - 2)^2 - (x_2 - 2)^2 + 9, \\ g_2(x) &= x_1^2 + x_2^2 - 4, \end{aligned}$$

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: The Frank-Wolfe method can be used to solve this optimization problem.

(1p) b) Let $f : \mathbb{R}^n \to \mathbb{R}$. Assume that f is twice continuously differentiable ($f \in C^2(\mathbb{R}^n)$) and strictly convex.

Claim: f is bounded from below.

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

$$f(x) = x_1^2 + \sin(x_1)x_2 - e^{x_2}.$$

Claim: The point $([0,\ 0]^T,\ 1)$ belongs to the epigraph of f, that is, $([0,\ 0]^T,\ 1) \in \operatorname{epi} f \subset \mathbb{R}^2 \times \mathbb{R}$.

(3p) Question 3

(Farkas' Lemma)

Farkas' Lemma is an important result in convex analysis and optimization. It can be stated as follows.

THEOREM: *Let* $A \in \mathbb{R}^{m\times n}$ *and* $b \in \mathbb{R}^m$*. Then, exactly one of the systems*

$$\begin{aligned} Ax &= b, \\ x &\geq 0, \end{aligned}$$

and

$$\begin{aligned} A^T y &\leq 0, \\ b^T y &> 0, \end{aligned}$$

has a feasible solution, and the other system is inconsistent.

Prove Farkas' Lemma. 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.

Question 4

(The Karush-Kuhn-Tucker conditions and Relaxations)

Consider the optimization problem

$$(1) \qquad \begin{cases} \underset{x\in\mathbb{R}^3}{\text{minimize}} & f(x), \\ \text{subject to} & g(x) \leq 0, \\ & h(x) = 0, \end{cases}$$

where

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

(1p) a) State the Karush-Kuhn-Tucker (KKT) conditions for the problem, and verify that $x^* = [-1/\sqrt{2}, -1/\sqrt{2}, 0]^T$ is a KKT point, that is, a feasible point for which there is a solution to the KKT-system.

(0.5p) b) Let $S \subset \mathbb{R}^n$ and $\tilde{f} : S \to \mathbb{R}$, and let $S_R \subset \mathbb{R}^n$ and $\tilde{f}_R : S_R \to \mathbb{R}$. Consider the two problems

$$(\text{P}) \qquad \begin{cases} \text{minimize} & \tilde{f}(x), \\ \text{subject to} & x \in S, \end{cases}$$

and

$$(\text{P}_\text{R}) \qquad \begin{cases} \text{minimize} & \tilde{f}_R(x), \\ \text{subject to} & x \in S_R. \end{cases}$$

Under what conditions do we call (P_R) a relaxation of (P)?

(1.5p) c) Use the results in a) and b) to show that x^* is globally optimal to (1). Motivate all steps and conclusions carefully.

(3p) Question 5

(Nonlinear programming)

Let N be a positive integer, and let $a^{(i)} > 0$ for $i = 1, 2, \dots, N$ be given positive real numbers. We define the geometric mean value of these positive real numbers as

$$\bar{a} = (a_1 a_2 \dots a_N)^{1/N} = \sqrt[N]{a_1 a_2 \dots a_N},$$

where for $a > 0$, $\sqrt[N]{a} = a^{1/N}$ denotes the Nth root. Show that $\bar{a}$ is a local minimizer of the problem

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\sum_{i=1}^{N}(\log(x) - \log(a_i))^2 \\ \text{subject to} \quad & x > 0, \end{aligned}$$

where log denotes the natural logarithm, i.e., $\log(e^x) = x$.

Hint 1: Recall the derivative of the natural logarithm, namely that for $x > 0$

$$\frac{d}{dx}\log(x) = \frac{1}{x}.$$

Hint 2: Recall the logarithm laws:

i) for $x, y > 0$, $\log(xy) = \log(x) + \log(y)$;

ii) for $p > 0$ and $x > 0$, $\log(x^p) = p\log(x)$.

(3p) Question 6

(Newton's method)

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by $f(x) = \frac{1}{2}x^T Q x$ where

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix},$$

and consider the problem

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

Starting in the point $x^{(0)} = [1, 1, 1]^T$ and using step length $\alpha = 1/2$, take one step in Newton's method with the Levenberg-Marquardt modification. As modification parameter γ, take the smallest valid integer for the modification.

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 7

(Modelling)

A friend of yours, we can call the person AR, is running a small business. The company is doing well, and AR would like to expand the business over the coming years. However, at the same time, AR would like to maximize the total revenue from the business over the same time horizon. AR has therefore asked for your help to derive a model for how the company should invest its yearly profit in order to maximize the total accumulated revenue over the given time period.

More specifically, the planning horizon is a period of T years. At the start of year 1, the company has assets (for example: buildings, machines, computers, etc.) of a total value of X SEK. Since AR is interested in maximizing the total accumulated revenue over the given time period, we model the total accumulated revenue at the beginning of year 1 to be 0 SEK. Based on historical data, a good model for the total profit that the company makes in one year is that it is a fraction α, where α is a fixed number in the interval $(0, 1)$, of that value of the companies assets at the beginning of the year.

AR only wants to take one investment decision per year. Therefore, you are asked to model it as if the company gets access to the previous years total profit at the start of the next year. At that time point, the company must choose what to do with the profit. It can do two things: It can add a fraction of the profit to the total accumulated revenue, and it can invest the remaining fraction of the profit in the company to grow the total value of the assets.

Finally, note that investments in the company to increase the assets cannot be taken out as revenue later on. This means that the total value of the companies assets cannot decrease over time. Moreover, it also means that the value of the company assets should not be counted as part of the total revenue over the time period. Nevertheless, since AR is not only interested in the total revenue, but is also interested in growing the company, a specific instruction was given to you: in your model, you must make sure that the company assets has at least doubled in size by the end of the planning period.

Help AR by formulating this as an optimization problem.

Hint: Consider using decision variables that model the state of relevant quantities, as well as actions taken, at the beginning of each year.

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 24–01–04
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

(Linear programming)

Transforming the problem to standard form gives

$$\begin{aligned} \text{minimize} \quad & -4x_1 - 2x_2 - 6x_3 \\ \text{subject to} \quad & 2x_1 - x_2 + 4x_3 + s = 30 \\ & x_1 + 2x_2 + 4x_3 = 80 \\ & x_1, x_2, x_3, s \geq 0. \end{aligned}$$

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

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

$x_B = B^{-1}b = [50/3,\, 35/3]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [-13/6,\, 1/3]$, and hence $(x_N)_1 = x_1$ enters the basis. $B^{-1}N_1 = [-1/3,\, 5/12]^T$, and since only the second component is positive we have that $(x_B)_2 = x_3$ leaves the basis.

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

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

$x_B = B^{-1}b = [28,\, 26]^T$. The reduced costs are $\tilde{c}_N^T = c_N^T - c_B^T B^{-1} N = [26/5,\, 6/5] > 0$. Hence the point $x^* = [28,\, 26,\, 0]^T$ is optimal.

Question 2

(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) False. The Frank-Wolfe method can only be applied when the constraint set is a polyhedron.

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

(1p) c) True. This is verified by $f([0,\, 0]^T) = 0 + 0 - 1 = -1 \leq 1$.

(3p) Question 3

(Farkas' Lemma)

See Theorem 4.35 or 10.10, and the corresponding proof.

Question 4

(the Karush-Kuhn-Tucker conditions and Relaxations)

(1p) a) For a feasible point x, i.e., a point such that $g(x) \leq 0$ and $h(x) = 0$, the KKT conditions are

$$\nabla f(x) + \mu \nabla g(x) + \lambda \nabla h(x) = 2 \begin{bmatrix} x_1 + 2 \\ x_2 + 2 \\ x_3 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 2x_1 \\ 2x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mu \geq 0$$

$$\mu g(x) = \mu(x_1 + x_2 + x_3) = 0.$$

It is easily verified that for the point $x^* = [-1/\sqrt{2}, -1/\sqrt{2}, 0]^T$, we have that $g(x^*) < 0$ and $h(x^*) = 0$, i.e., x^* is feasible and g is not active. Plugging the point into the KKT conditions, we find that equations are satisfied for $\mu^* = 0$ and $\lambda^* = 2\sqrt{2} - 1$.

(0.5p) b) The two conditions are $S \subseteq S_R$, and $f_R(x) \leq f(x)$ for $x \in S$.

(1.5p) c) If the constraint $h(x) = 0$ is changed to $h(x) \leq 0$, we get a relaxation of the original problem (why?). Moreover, the relaxed problem is convex (why?). Furthermore, since multiplier λ^* for the equality constraint is positive in the point $x^* = [-1/\sqrt{2}, -1/\sqrt{2}, 0]^T$, x^* remains a KKT point for the relaxed problem (why?). By sufficiency of KKT points for global optimality for convex problems (Theorem 5.49), x^* is globally optimal for the relaxed problem. By the relaxation theorem (Theorem 6.1), it is therefore also globally optimal for the original problem.

(3p) Question 5

(Nonlinear programming)

The optimization problem is a one-dimensional problem. Denoting the cost function $f(x)$, for $x > 0$ the derivative is given by

$$f'(x) = \frac{1}{2}\sum_{i=1}^{N}\left(2(\log(x) - \log(a_i))\frac{1}{x}\right) = \sum_{i=1}^{N}\frac{\log(x) - \log(a_i)}{x} = \frac{N\log(x) - \sum_{i=1}^{N}\log(a_i)}{x}.$$

Since the feasible region is open, and since f is differentiable on the feasible region, if a local minimizer $x^\star$ exists, it must fulfill $f'(x^\star) = 0$ (motivate!). This means that a potential local minimizer must fulfill

$$\frac{N\log(x^\star) - \sum_{i=1}^{N}\log(a_i)}{x^\star} = 0 \quad \iff \quad \log(x^\star) = \frac{1}{N}\sum_{i=1}^{N}\log(a_i).$$

Using the logarithm laws, this gives

$$\log(x^\star) = \frac{1}{N}\sum_{i=1}^{N}\log(a_i) = \frac{1}{N}\log(a_1a_2\ldots a_N) = \log\left((a_1a_2\ldots a_N)^{1/N}\right).$$

Since log is bijective, we have that $x^\star = \bar{a} = (a_1a_2\ldots a_N)^{1/N} > 0$.

For $x > 0$, the second derivative is given by

$$f''(x) = N\frac{\frac{1}{x}x - \log(x)}{x^2} + \frac{\sum_{i=1}^{N}\log(a_i)}{x^2} = \frac{N - N\log(x) + \sum_{i=1}^{N}\log(a_i)}{x^2}.$$

A direct caluclation gives that

$$f''(x^\star) = \frac{N}{(x^\star)^2} > 0,$$

and hence $x^\star$ is a local minimum (motivate!).

(3p) Question 6

(Newton's method)

We have that $\nabla f(x) = Qx$ and $\nabla^2 f(x) = Q$. The eigenvalues of Q are $\lambda_1 = 1$ and $\lambda_{2,3} = 1/2 \pm \sqrt{13/4}$. So there is one negative eigenvalue of the Hessian, and it is given by $1/2 - \sqrt{13/4} \approx -1.3028$. That means that $\gamma = 2$. The search direction is therefore given by

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

This means that

$$x^{(1)} = x^{(0)} + \alpha p^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -\frac{1}{3} \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ \frac{1}{2} \\ \frac{3}{2} \end{bmatrix}.$$

(3p) Question 7

(Modelling)

For each $t = 1, 2, \dots, T+1$, let

$$\begin{aligned} x_t &= \text{ total value of assets at the beginning of year } t \\ y_t &= \text{ total revenue accumulated at the beginning of year } t. \end{aligned}$$

This means that x_{T+1} and y_{T+1} are the total value of assets and total accumulated revenues, respectively, at the end of the planning horizon. For $t = 1, 2, \dots, T$, let

$$\gamma_t = \text{ fraction of profit from year } t \text{ invested in assets at the start of year } t+1$$

An optimization problem for maximizing total revenue accumulated over the time period is given by

$$\begin{aligned} \max_{x_t,\, y_t,\, \gamma_t} \quad & y_{T+1} \\ \text{subject to} \quad & x_{t+1} = x_t + \gamma_t \alpha x_t, \quad \forall t \in \{1, 2, \dots, T\} && \text{(accumulated value of assets)} \\ & x_1 = X, && \text{(starting value for assets)} \\ & y_{t+1} = y_t + (1 - \gamma_t) \alpha x_t, \quad \forall t \in \{1, 2, \dots, T\} && \text{(accumulated revenue)} \\ & y_1 = 0, && \text{(starting value for revenue)} \\ & x_{T+1} \geq 2X, && \text{(assets at least double in size)} \\ & 0 \leq \gamma_t \leq 1, \quad \forall t \in \{1, 2, \dots, T\}. \end{aligned}$$
