Chalmers/GU
Mathematics sciences

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	20–08-18
Time:	8^{30}–13^{30}
Aids:	All aids are allowed, but cooperation is not allowed
Number of questions:	7; passed on one question requires 2 points of 3. Questions are *not* numbered by difficulty. To pass requires 10 points and three passed questions.
Examiner:	Ann-Brith Strömberg
Note:	It is not possible to "plus" (retaking an exam in a course you have already passed, to raise its grade). Students who have not yet passed the exam can attend this re-exam.

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

(Simplex method) Consider the problem

$$\begin{aligned} \text{maximize} \quad & x_1 - x_2 \\ \text{subject to} \quad & 2x_1 + x_2 \geq 2 \\ & x_1 - x_2 \leq 2 \\ & x_1, \quad x_2 \geq 0 \end{aligned}$$

(0.5p) a) Convert the problem to standard form.

(1.5p) b) Solve the problem using Phase I and Phase II of the simplex method. Use you calculations to provide an optimal solution or a unbounded ray in the original variables.

(1p) c) Derive the set of optimal solutions by analysing the reduced costs of the final iteration and conducting another minimum ratio test.

Question 2

(Representation theorem)

Consider the problem to $\text{minimize}_{\boldsymbol{x} \in P} f(\boldsymbol{x})$, where P is a non-empty polyhedron.

(2p) a) Assume that f is a concave function and that the problem has an optimal solution. Does the set of optimal solutions contain an extreme point of P? Prove or provide a counter example.

(1p) b) Assume that f is a convex function. Does the set of optimal solutions always contain an extreme point of P? Prove or provide a counter example.

(3p) Question 3

(Convexity)

Let $f_1, f_2, \dots, f_k : \mathbb{R}^n \to \mathbb{R}$ be convex functions. Consider the function f defined by $f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \dots, f_k(\boldsymbol{x})\}$.

(2p) a) Show that f is convex.

(1p) b) State and prove a similar result for concave functions.

(3p) Question 4

(Linear programming)

Consider the problem to

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned} \tag{1}$$

and the perturbed version of the problem where the right-hand-side is changed from $\boldsymbol{b}$ to $\boldsymbol{b} + \delta\boldsymbol{b}$. Show that if the original problem (1) has an optimal solution then the perturbed version cannot be unbounded, independently of $\delta\boldsymbol{b}$.

(3p) Question 5

(Modelling)

A rocket launching problem. Suppose that we are to send a rocket to the altitude of $\bar{z}$ [m] in time T [s]. Let $z(t)$ [m] denote its height above the ground at time t and $f(t)$ [N] be the non-negative upward force of the rocket thrusters at time t. Let the mass of the rocket be m [kg], the maximal thrust of the rocket be b [N], an let $v(t) = z'(t)$ [m/s] denote the upward velocity.

Formulate an optimization problem, with a quadratic objective function and affine constraints, that minimizes the energy required for the rocket to reach the desired altitude at time T.

Hints: The amount of energy required can be computed by

$$\int_0^T f(t)v(t)\,dt,$$

and the equation of motion is

$$mv'(t) + mg = f(t), \quad t \in [0, T].$$

Assume that the time interval is divided into K periods of length $l := T/K$ and let $f_k := f(lk)$, $z_k := z(lk)$, $v_k := v(lk)$, $k = 1, \dots K$. Then approximate the velocity and acceleration using finite differences, e.g., $v_k = (z_k - z_{k-1})/l$, $k = 1, \dots, K$. Similarly, approximate the integral as a Riemann sum.

Question 6

(true or false)

Indicate for each of the following three statements whether it is true or false. Motivate your answers!

(1p) a) *Claim:* The Simplex method is a suitable solution method for problems where a convex objective function should be optimized over a polytope.

(1p) b) *Claim:* For a convex optimization problem, every KKT-point is a global optimal solution.

(1p) c) Consider a convex function $f : \mathbb{R}^n \mapsto \mathbb{R}$.

Claim: If f is differentiable at a point $\bar{\boldsymbol{x}} \in \mathbb{R}^n$, then the identity $\partial f(\bar{\boldsymbol{x}}) = \{\nabla f(\bar{\boldsymbol{x}})\}$ holds.

(3p) Question 7

(Exterior penalty method)

Consider the following problem:

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := 2e^{x_1} + 3x_1^2 + 2x_1x_2 + 4x_2^2, \\ \text{subject to} \quad & 3x_1 + 2x_2 - 6 = 0. \end{aligned}$$

Formulate a suitable exterior penalty function with the penalty parameter $\nu = 10$. Starting at the point $(1, 1)$, perform one iteration of a gradient method to solve the unconstrained penalty problem.

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 20–08–18
Examiner: Ann-Brith Strömberg

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

Question 1

(Simplex method)

(0.5p) a) The problem on standard form is:

$$\begin{aligned} \text{minimize } & -x_1 + x_2 \\ \text{subject to } & 2x_1 + x_2 - s_1 = 2 \\ & x_1 - x_2 + s_2 = 2 \\ & x_1, \quad x_2, \quad s_1, \quad s_2 \geq 0 \end{aligned}$$

(1.5p) b) Utilizing that s_2 can be for the initial BFS, the phase I problem is

$$\begin{aligned} \text{minimize } & + a_1 \\ \text{subject to } & 2x_1 + x_2 - s_1 + a_1 = 2 \\ & x_1 - x_2 + s_2 = 2 \\ & x_1, \quad x_2, \quad s_1, \quad s_2, \quad a_1 \geq 0 \end{aligned}$$

Our basic variables are (a_1, s_2) and our non-basic are (x_1, x_2, s_1), we get

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

The reduced costs for the non-basic variables are

$$\tilde{c}_N^T = c_N^T - \tilde{c}_B^T B^{-1} N = -\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & -1 & 1 \end{bmatrix},$$

by the minimum reduced cost rule, x_1 enter the basis. We have that $B^{-1}N_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, the minimum ratio test is thus

$$\underset{i|(B^{-1}N_1)_i>0}{\operatorname{argmin}} \frac{(x_B)_i}{(B^{-1}N_1)_i} = \underset{i|(B^{-1}N_1)_i>0}{\operatorname{argmin}} \begin{bmatrix} \frac{2}{2} & \frac{2}{1} \end{bmatrix}$$

And thus a_1 leaves the basis and Phase I is complete.

Our basic variables are (x_1, s_2) and our non-basic are (x_2, s_1), we get

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

The reduced costs for the non-basic variables are

$$\tilde{c}_N^T = \begin{bmatrix} 1 & 0 \end{bmatrix} - \underbrace{\begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}}_{=\begin{bmatrix} -\frac{1}{2} & 0 \end{bmatrix}} \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \end{bmatrix},$$

by the minimum reduced cost rule, s_1 enter the basis. We have that $B^{-1}N_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$, the only positive denominator in the minimum ratio test corresponds to s_2, which leaves the basis.

Our basic variables are (x_1, s_1) and our non-basic are (x_2, s_2), we get

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

The reduced costs for the non-basic variables are

$$\bar{c}_N^T = \begin{bmatrix} 1 & 0 \end{bmatrix} - \underbrace{\begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}}_{=\begin{bmatrix} 0 & -1 \end{bmatrix}} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \end{bmatrix} \geq 0,$$

since the reduced costs are all non-negative, we conclude that the current basis is optimal, and the values of the original variables are $\boldsymbol{x} = (2, 0)$.

(1p) c) Since the reduced costs of s_2 is strictly positive we deduce that $s_2^* = 0$. We let x_2 enter the basis and do the minimum ratio test. Note that $B^{-1}N_1 = \begin{bmatrix} -1 \\ -3 \end{bmatrix}$ imply that the entire ray

$$x_B = B^{-1}b - \gamma B^{-1}N_1,\ x_2 = \gamma,\ s_2 = 0,\ \gamma \geq 0,$$

is feasible. Since the reduced costs of x_2 is zero we yield that the ray is a set of optimal solutions. Returning to the original variables we get that $\boldsymbol{x} = (2 + \gamma, \gamma)$ is an optimal solution for each $\gamma \geq 0$. Noting that this is precisely the set for which $s_2 = 0$ and thus it equals the set of optimal solutions.

Question 2

(Representation theorem)

(2p) a) Let x_i, $i \in I$ be the extreme points and d_j, $j \in J$ be the extreme directions of P, respectively. Then we have by the representation theorem that

$$P = \left\{ \sum_{i \in I} \lambda_i x_i + \sum_{j \in J} \mu_j d_j \;\middle|\; \sum_{i \in I} \lambda_i = 1, \boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0 \right\}.$$

Now, consider the optimal solution $x^* \in P$ that exists by assumption, i..e, $f(x^*) \leq f(x), x \in P$.

First we will show that $\mu^* = 0$ or that such a choice exists. Let $j \in J$ be such that $\mu_j^* > 0$ and consider the line segment between $\mu_j^1 = 0, \mu_j^2 = 2\mu_j^*$, and let x^1, x^2 be the corresponding points, by the concavity of f we have that $f(x^1)/2 + f(x^2)/2 \leq f(x^*)$. Hence, by the optimality of x^* we yield that $f(x^1) = f(x^2) = f(x^*)$, showing that $\mu_j = 0$ is also a optimal choice.

Similarly assume that x^* is an optimal solution but not an extreme point. By the concavity of f we have that

$$f(x^*) = f(\sum_{j \in I} \lambda_i x_i) \geq \sum_{j \in I} \lambda_i f(x_i)$$

However, since $f(x_i) \geq f(x^*)$, we get that $\lambda_i = 0$ if $f(x_i) > f(x^*)$ and for $\lambda_i > 0$, $f(x^i) = f(x^*)$. Thus, x^* is a convex combination of optimal extreme points.

(1p) b) Consider the counter-example, $f(x) = x^2, P = [-1, 1]$, the extreme-points are clearly non-optimal.

(3p) Question 3

(Convexity)

(1.5p) a) Consider $\bar{x} = \lambda x^1 + (1-\lambda)x^2$.

$$\begin{aligned} f(\bar{x}) &= \max\{f_1(\bar{x}), f_2(\bar{x}), \ldots, f_k(\bar{x})\} \\ &\quad \text{since } f_i(x) \text{ convex} \\ &\le \max\{\lambda f_1(x^1) + (1-\lambda)f_1(x^2), \ldots, \lambda f_k(x^1) + (1-\lambda)f_k(x^2)\} \\ &\le \lambda \max\{f_1(x^1), \ldots, f_k(x^1)\} + (1-\lambda)\max\{f_1(x^2), \ldots, f_k(x^2)\} \\ &= \lambda f(x^1) + (1-\lambda)f(x^2) \end{aligned}$$

By the definition of a convex function, f is convex.

(1.5p) b) Let $g_1, g_2, \ldots, g_k : R^n \to R$ be concave functions. Consider the function g defined by $g(\mathbf{x}) = \min\{g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_k(\mathbf{x})\}$. g is a concave function.

Proof: Set $\bar{f}_1 = -g_1, \ldots, \bar{f}_k = -g_k$. We get $\bar{f} = -g$. Since $g_1, g_2, \ldots, g_k$ are concave functions, $f_1, f_2, \ldots, f_k$ are convex functions. From above, we know f is convex, so g is concave.

(3p) Question 4

(Linear programming) Use Strong duality to realize that the dual problem to (1) also must have an optimal solution, and hence, a feasible solution.

This feasibility does not change if $\boldsymbol{b}$ is perturbed to $\boldsymbol{b}+\delta\boldsymbol{b}$, independently of $\delta\boldsymbol{b}$. Which, by using Weak duality, implies that the perturbed problem cannot be unbounded.

(3p) Question 5

(modeling) Using the variables and parameters introduced in the question but extending to also include v_0 and z_0, we yield that the problem is to

$$\begin{aligned} &\text{minimize} && l\sum_{k=1}^{K} f_k v_k && && (1) \\ &\text{subject to} && z_k - z_{k-1} = l v_k, && k = 1, \ldots, K && (2) \\ & && \frac{m}{l}(v_k - v_{k-1}) = f_k - mg, && k = 1, \ldots, K && (3) \\ & && f_k \le b, && k = 1, \ldots, K && (4) \\ & && f_k, z_k \ge 0, && k = 1, \ldots, K && (5) \\ & && z_K = \bar{z} && && (6) \\ & && v_0 = z_0 = 0 && && (7) \end{aligned}$$

Question 6

(true or false)

(1p) a) False. The Simplex method is used for linear optimization problems.

(1p) b) True. See theorem regarding sufficiency of the KKT conditions for convex optimization problems in the textbook.

(1p) c) True. See theorem in the textbook regarding subgradients.

(3p) Question 7

(Exterior penalty method)

Using the quadratic penalty function, the penalty problem is given as follows:

$$\text{minimize} \quad F_\nu(\boldsymbol{x}) = 2e^{x_1} + 3x_1^2 + 2x_1x_2 + 4x_2^2 + \nu[3x_1 + 2x_2 - 6]^2$$

$$\nabla F_\nu(\boldsymbol{x}) = \begin{bmatrix} 2e^{x_1} + 6x_1 + 2x_2 + 6\nu[3x_1 + 2x_2 - 6] \\ 2x_1 + 8x_2 + 4\nu[3x_1 + 2x_2 - 6] \end{bmatrix}$$

Since the penalty parameter $\nu = 10$, we get

$$F_\nu(\boldsymbol{x}) = 2e^{x_1} + 3x_1^2 + 2x_1x_2 + 4x_2^2 + 10[3x_1 + 2x_2 - 6]^2$$

$$\nabla F_\nu(\boldsymbol{x}) = \begin{bmatrix} 2e^{x_1} + 186x_1 + 122x_2 - 360 \\ 122x_1 + 88x_2 - 240 \end{bmatrix}$$

Apply steepest descent method with exact line search,

$$\boldsymbol{x}^1 = (1,1)^T,\ \nabla F_\nu(\boldsymbol{x}) = \begin{bmatrix} 2e - 52 \\ -30 \end{bmatrix},\ d^1 = -\nabla F_\nu(\boldsymbol{x}) = \begin{bmatrix} 52 - 2e \\ 30 \end{bmatrix}.$$

Solve the minimization problem min $F_\nu(\boldsymbol{x}^1 + \lambda d^1)$, we get the step length $\lambda^* = 0.004$, so

$$\boldsymbol{x}^2 = \boldsymbol{x}^1 + \lambda^* d^1 = [1.86, 1.12]^T$$