Chalmers/Gothenburg University
Mathematical Sciences

EXAM

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date:	08–12–15
Time:	House V, morning
Aids:	Text memory-less calculator, English–Swedish dictionary
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:	Michael Patriksson
Teacher on duty:	Peter Lindroth (0762-721860)
Result announced:	09–01–16 Short answers are also given at the end of the exam on the notice board for optimization in the MV building.

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.

At the end of the exam

Sort your solutions by the order of the questions.
Mark on the cover the questions you have answered.
Count the number of sheets you hand in and fill in the number on the cover.

Question 1

(the simplex method)

Consider the following linear program:

$$
\begin{array}{llrl}
\text{minimize} & z = & -x_1 + \ \ x_2, & \\
\text{subject to} & & -x_1 + 2x_2 & \geq 1/2, \\
& & -2x_1 - 2x_2 & \geq 1, \\
& & x_1 \qquad\quad & \in \mathbb{R} \ \text{(free)}, \\
& & x_2 & \geq 0.
\end{array}
$$

(2p) a) Solve this problem by using phase I and phase II of the simplex method.

[*Aid:* Utilize the identity

$$
\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}
$$

for producing basis inverses.]

(1p) b) Without solving the dual to the problem above, motivate clearly whether there are no optimal dual solutions, a unique optimal dual solution (if so, present it) or multiple optimal dual solutions (if so, present at least two of them).

(3p) **Question 2**

(convergence of an exterior penalty method)

Let us consider a general optimization problem:

$$\begin{aligned} &\text{minimize } f(\boldsymbol{x}), \\ &\text{subject to } \boldsymbol{x} \in S, \end{aligned} \tag{1}$$

where $S \subset \mathbb{R}^n$ is a non-empty, closed set and $f : \mathbb{R}^n \to \mathbb{R}$ is a given differentiable function. We assume that the feasible set S of the optimization problem (1) is given by the system of inequality and equality constraints:

$$\begin{aligned} S = \{\, \boldsymbol{x} \in \mathbb{R}^n \mid g_i(\boldsymbol{x}) &\leq 0, \quad i = 1, \dots, m, \\ h_j(\boldsymbol{x}) &= 0, \quad j = 1, \dots, \ell \,\}, \end{aligned} \tag{2}$$

where $g_i \in C(\mathbb{R}^n)$, $i = 1, \dots, m$, $h_j \in C(\mathbb{R}^n)$, $j = 1, \dots, \ell$.

We choose a function $\psi : \mathbb{R} \to \mathbb{R}_+$ such that $\psi(s) = 0$ if and only if $s = 0$ (typical examples of $\psi(\cdot)$ are $\psi_1(s) = |s|$, or $\psi_2(s) = s^2$), and introduce the function

$$\nu \check{\chi}_S(\boldsymbol{x}) := \nu \Bigg(\sum_{i=1}^{m} \psi\big(\max\{0, g_i(\boldsymbol{x})\}\big) + \sum_{j=1}^{\ell} \psi\big(h_j(\boldsymbol{x})\big) \Bigg), \tag{3}$$

where the real number $\nu > 0$ is called a *penalty parameter*.

We assume that for every $\nu > 0$ the approximating optimization problem to

$$\text{minimize } f(\boldsymbol{x}) + \nu \check{\chi}_S(\boldsymbol{x}) \tag{4}$$

has at least one optimal solution $\boldsymbol{x}^*_\nu$.

We then have the following result.

Theorem 1 *Assume that the original constrained problem* (1) *possesses optimal solutions. Then, every limit point of the sequence* $\{\boldsymbol{x}^*_\nu\}$, $\nu \to +\infty$, *of globally optimal solutions to* (4) *is globally optimal in the problem* (1).

Prove this theorem.

(3p) **Question 3**

(applications of Weierstrass' Theorem)

For each of the following functions f_i, $i = 1, 2, 3$, motivate carefully if a global minimum is attained on the corresponding set S_i, $i = 1, 2, 3$.

(1) $f_1(\boldsymbol{x}) = -e^{-\frac{(x_1+2)^2+(x_2+1)^2}{10}} + 10e^{-\frac{(x_1+2)^2+(x_2+1)^2}{100}} + \frac{1}{50}\big((x_1+2)^2 + (x_2+1)^2\big) + \frac{1}{10}x_1,$
$S_1 = \mathbb{R}^2.$

(2) $f_2(\boldsymbol{x}) = \begin{cases} -\frac{1}{x_1^2+(x_2-1)^2+2x_3^2} + x_3^2, & \text{if } x_1 > 0, \\ 0, & \text{if } x_1 \leq 0, \end{cases}$
$S_2 = \left\{ \boldsymbol{x} \in \mathbb{R}^3 \mid -5 \leq x_i \leq 5,\ \forall i \right\}.$

(3) $f_3(\boldsymbol{x}) = \left(x_1 + x_2^2\right)^2 + x_1 + 3x_2 + 200,$
$S_3 = \mathbb{R}^2.$

Question 4

(modeling)

The government has assigned you to lead their aid program. They are willing to spend 1 % of the national gross product of b SEK. They are considering to give aid to a set of countries $\mathcal{N} = \{1, \dots, N\}$. The aim of the aid program is to increase the Human Development Index (HDI) in the countries, which is calculated by measuring the three factors education per capita, life expectancy and gross national product per capita (GDP). We may therefore consider the HDI index as a measure of *development per capita* in a country.

For all countries $j \in \mathcal{N}$ let a_j denote the current value of the HDI index, c_j the increase of HDI per SEK given as aid to the country and p_j the population size. There are only a limited number of aid programs in each of the countries. This puts a limit on the maximal aid that a country can receive, let d_j SEK denote the maximal aid country j can receive.

(1p) a) Write a linear program for distributing aid that maximizes the total HDI in the region formed by all the countries considered.

(1p) b) You are given new directives from the government: they think that the aid should be focused on a maximal number of M countries. Introduce integer variables in your model (i.e., write an integer linear programming model) in order to accommodate this demand.

(1p) c) There has recently been some discussions concerning the aid to countries with a high HDI. The government wants you to write a new model that maximizes the minimal HDI among the countries. Extend your model in a) to an LP model that accommodates this demand.

(3p) Question 5

(the Frank–Wolfe method)

Consider the optimization problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := x_1^2 + x_1x_2 + 2x_2^2 - 10x_1 - 4x_2, \\ \text{subject to} \quad & x_1 + x_2 \le 3, \\ & 0 \le x_1 \le 2, \quad 0 \le x_2 \le 2. \end{aligned}$$

Start at the point $\boldsymbol{x}^0 = (0,0)^{\mathrm{T}}$ and perform two iterations of the Frank–Wolfe method. (Recall that the Frank–Wolfe method starts at some feasible point. Given an iteration k and feasible iterate $\boldsymbol{x}_k$ it produces a feasible search direction $\boldsymbol{p}_k$ through the minimization of the first-order Taylor expansson of f at $\boldsymbol{x}_k$. The next iterate is found through an exact line search in f along the search direction, such that the resulting vector is also feasible.) Write out the upper and lower bounds for the optimal objective function value that the algorithm generates in each iteration, and give a theoretical motivation to them. If an optimum is found, state so, and motivate why it is an optimum.

(3p) Question 6

(convex problem)

Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := -\ln(x_1 + x_2) + x_3 \ln x_3, \\ \text{subject to} \quad & g_i(\boldsymbol{x}) := -x_i + 1 \le 0, \qquad i = 1,2,3, \\ & g_4(\boldsymbol{x}) := -x_1 + 2x_2^2 + 4x_3^2 - 10 \le 0. \end{aligned}$$

Establish whether this is a convex problem or not.

[*Note:* By "convex problem" we refer to the property that the objective function is convex in a minimization problem, and that the feasible set is a convex set.]

(3p) Question 7

(linear programming duality)

Consider the following two polyhedral sets corresponding to the feasible sets of the standard pair of primal–dual linear programs:

$$X = \{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}, \quad \boldsymbol{x} \geq \boldsymbol{0}^n \},$$
$$Y = \{ \boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c}, \quad \boldsymbol{y} \geq \boldsymbol{0}^m \}.$$

Prove that if both X and Y are non-empty, then at least one of them must be unbounded.

[*Remark:* This result can in fact be strengthened to the following: If at least one of the sets X and Y is non-empty, then at least one of them is non-empty and unbounded; this result is due to Clark (1961).]

Good luck!

Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 08–12–15

Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) To transform the problem to standard form, the free variable x_1 must be replaced by the non-negative variables x_1^+ and x_1^- such that $x_1 := x_1^+ - x_1^-$. A non-negative slack variable s_1 in the first constraint and a non-negative slack variable s_2 in the second constrained must be subtracted.

A BFS cannot be found directly, hence begin with phase 1 with artificial variables $a_1 \geq 0$ added in the first constraint and $a_2 \geq 0$ in the second constraint. The objective is to minimize $w = a_1 + a_2$. Start with the BFS given by $(a_1, a_2)^{\mathrm{T}}$. In the first iteration of the simplex algorithm, x_1^- is the only variable with a negative reduced cost (-3), and is therefore the only eligable incoming variable. The minimum ratio test shows that either a_1 or a_2 can be removed from the basis. We choose a_1 as the outgoing variable and update the basic variables to $\boldsymbol{x}_B = (x_1^-, a_2)^{\mathrm{T}}$. By computing the reduced costs, we see that s_1 is the only non-basic variable with negative reduced cost (-2) and s_1 is chosen as incoming variable. The minimum ratio test shows that a_2 should leave the basis. By updating the basis and computing the reduced costs we see that we are now optimal with $w^* = 0$ and we proceed to phase 2.

The BFS is given by $\boldsymbol{x}_B = (x_1^-, s_1)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_1^+, x_2, s_2)^{\mathrm{T}}$ and the reduced costs with the phase 2 cost vector $\boldsymbol{c} = (-1, 1, 1, -1, -1)^{\mathrm{T}}$ are

$$\tilde{\boldsymbol{c}}^{\mathrm{T}}_{(x_1^+, x_2, s_2)} = (0, 2, 1/2) \geq \boldsymbol{0}^3,$$

and thus the optimality condition is fulfilled for the current basis. We have $\boldsymbol{x}_B^* = (1/2, 0)^{\mathrm{T}}$, or in the original variables, $\boldsymbol{x}^* = (x_1, x_2)^* = (-1/2, 0)^{\mathrm{T}}$, with the optimal value $z^* = 1/2$.

(1p) b) Since there is an optimal solution to the problem, Strong duality guarantees the existence of a dual optimal solution. The expression for this is $\boldsymbol{y}^{*\mathrm{T}} = \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} = (0, 1/2)$. However, the optimal basis is degenerate and it is possible to replace the zero-valued basic variable s_2 with a non-basic variable as long as the basis matrix $\boldsymbol{B}$ still has linear independent columns. We see that it is possible to replace s_1 with s_2 which gives us $\boldsymbol{y}^{*\mathrm{T}} = \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} = (2/3, 1/6)$ or to replace s_1 with x_2 which gives us $\boldsymbol{y}^{*\mathrm{T}} = \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} = (1, 0)$. Also, any convex combination between these three dual optimal solution is also dual optimal.

(3p) Question 2

(convergence of an exterior penalty method)

See Theorem 13.3 in The Book.

Question 3

(Weierstrass)

(1p) a) Yes it does. The function is continuous on a closed set. Also, we observe that the function is weakly coercice, i.e., when $||\boldsymbol{x}|| \to \infty$, then $f_1(\boldsymbol{x}) \to \infty$. Weierstrass' theorem now guarantees that a global minimum exists.

(1p) b) No it does not. The function is not lower semi-continuous, so we cannot invoke Weierstrass' thereom. We observe that along the arc $\boldsymbol{x} = (t, 1, 0)^{\mathrm{T}}$, where $t \to 0_+$, $f_2(\boldsymbol{x}) \to -\infty$.

(1p) c) No it does not. The function is not weakly coercive, so we cannot invoke Weierstrass theorem. We observe that along the arc $\boldsymbol{x} = (t, -\sqrt{-t})^{\mathrm{T}}$, where $t \to -\infty$, $f_3(\boldsymbol{x}) \to -\infty$.

Question 4

(modeling)

(1p) a) Introduce the variables:
x_i : SEK aid given to country i.

The objective is to

$$\max \quad \frac{\sum_{i\in\mathcal{N}}(a_i + c_i x_i)p_i}{\sum_{i\in\mathcal{N}} p_i},$$

and the constraints are

$$\sum_{i\in\mathcal{N}} x_i \leq 0.01b, \tag{1}$$

$$0 \leq x_i \leq d_i, \qquad \forall i \in \mathcal{N}. \tag{2}$$

(1p) b) Introduce the binary variables
y_i : with value one if country i receives aid, zero otherwise.

Modify the constraints (2) into

$$0 \leq x_i \leq d_i y_i, \quad \forall i \in \mathcal{N}. \tag{3}$$

Introduce an additional constraint

$$\sum_{i \in \mathcal{N}} y_i \leq M. \tag{4}$$

(1p) c) Introduce the variable
w : minimal HDI of the countries considered.

Introduce the constraints

$$a_i + c_i x_i \geq w \quad \forall i \in \mathcal{N}. \tag{5}$$

Change the objective function into

$$\max w.$$

(3p) Question 5

(the Frank–Wolfe method)

Iteration 1: $\boldsymbol{x}_0 = (0,0)^{\mathrm{T}}$, $f(\boldsymbol{x}_0) = 0$. It is feasible, so we get an upper bound: $[LBD, UBD] = (-\inf, 0]$. We have that $\nabla f(\boldsymbol{x}_0) = (-10, -4)^{\mathrm{T}}$. Solve the LP $\min_{y \in X} \nabla f(\boldsymbol{x}_0)^{\mathrm{T}} \boldsymbol{y}$. The solution is obtained at $\boldsymbol{y}_0 = (2,1)^{\mathrm{T}}$. The search direction is $\boldsymbol{p}_0 = \boldsymbol{y}_0 - \boldsymbol{x}_0 = (2,1)^{\mathrm{T}}$. Since f is convex, $g(\boldsymbol{y}) := f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}_0) \leq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathbb{R}^2$. The LP problem $\min_{y \in X} g(\boldsymbol{y})$ is a relaxation of the original problem, hence an optimal objective value gives a lower bound. The objective value is $f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^{\mathrm{T}}(\boldsymbol{y}_0 - \boldsymbol{x}_0) = 0 + (-10, -4)(2,1)^{\mathrm{T}} = -24$. Hence $[LBD, UBD] = [-24, 0]$. Line search: $\phi(\alpha) = f(x_0 + \alpha \boldsymbol{p}_0) = f((2\alpha, \alpha)) = \ldots = 8\alpha^2 - 24\alpha$. $\phi'(\alpha) = 16\alpha - 24 = 0$. $\alpha = 24/16 > 1$ Take a unit step: $\alpha = 1$. Hence, the next point is $\boldsymbol{x}_1 = (2,1)^{\mathrm{T}}$.

Iteration 2: $f(\boldsymbol{x}_1) = -16$. So $[LBD, UBD] = [-24, -16]$. $\nabla f(\boldsymbol{x}_1) = (-5, 2)^{\mathrm{T}}$, the optimal solution to $\min_{y \in X} \nabla f(\boldsymbol{x}_1)^{\mathrm{T}} \boldsymbol{y}$ is obtained at $\boldsymbol{y}_1 = (2,0)^{\mathrm{T}}$, hence the search direction will be $\boldsymbol{p}_1 = (\boldsymbol{y}_1 - \boldsymbol{x}_1) = (0, -1)^{\mathrm{T}}$. Since $f(\boldsymbol{x}_1) + \nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{y}_1 - \boldsymbol{x}_1) = -18$, we have $[LBD, UBD] = [-18, -16]$. Line search: $\phi(\alpha) = f(\boldsymbol{x}_1 + \alpha \boldsymbol{p}_1) = 2^2 + 2(1-\alpha) + 2(1-\alpha)^2 - 20 - 4(1-\alpha)$. $\phi'(\alpha) = -2 - 4(1-\alpha) + 4 = 0$. $\alpha = 1/2$. $\boldsymbol{x}_3 = (2, 1/2)^{\mathrm{T}}$.

The point $\boldsymbol{x}_3$ is a KKT point since $\nabla f(\boldsymbol{x}_3) = (-11/2, 0)$ and the active constraint is $g(\boldsymbol{x}) = x_1 - 2$ with $\nabla g(\boldsymbol{x}_3) = (1, 0)^{\mathrm{T}}$. The objective function is convex (eigenvalues of the Hessian are all non-negative) and the feasible region is a polyhedron, so the problem is convex. A KKT point is sufficient for optimality in convex problems, and $\boldsymbol{x}_3$ is therefore an optimal point.

(3p) Question 6

(convex problem)

We first conclude that the feasible set is convex. The functions g_i, $i = 1, 2, 3$, are affine, hence convex. The function g_4 is convex, since its Hessian matrix is constant and diagonal with diagonal entries 0, 4, and 8, which all are non-negative. In each of these four cases, the constraint is of the form $g_i(\boldsymbol{x}) \leq 0$; hence, by Proposition 3.44, each feasible set is convex, and moreover their intersection is convex by Proposition 3.3.

To establish that the objective function is convex on the convex feasible set of the problem at hand, we consider the function terms one by one. The function $\boldsymbol{x} \mapsto -\ln(x_1 + x_2)$ is of the form $-\ln t$, where $t = x_1 + x_2$. Introducing, for simplicity, t as an additional variable, we notice that the equation just given is linear and therefore represent a further convex constraint. Due to constraints 1 and 2, $t > 0$ on the feasible set, whence ln is well defined there. Finally, $-\ln t$ is a (strictly) convex function on this domain. The second term of the objective is $x_3 \ln x_3$. Again, we get from the third constraint that $x_3 > 0$, and hence the term is well-defined. Taking its derivative with respect to x_3 we get $1 + \ln x_3$, and its derivative is, in turn, $1/x_3$, which is positive. Hence, the objective function is a sum of two convex functions and therefore is convex. We are done.

(3p) Question 7

(linear programming duality)

Suppose, for example, that X is bounded. Then, there exists a bounded optimal solution for every value of the objective coefficient vector $\boldsymbol{c}$. Therefore, its dual must also have bounded optimal solutions for every value of $\boldsymbol{c}$. It follows that the dual problem must have feasible solutions for every $\boldsymbol{c}$. Consider the cone

$$C := \{ \boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \leq \mathbf{0}^n, \quad \boldsymbol{y} \geq \mathbf{0}^m \}.$$

By the Representation Theorem, the set Y is bounded if and only if C contains only the zero vector. By the above, the set $\{ \boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \leq -\boldsymbol{e}, \quad \boldsymbol{y} \geq \mathbf{0}^m \}$, where $\boldsymbol{e}$ is the m-vector of ones, is non-empty. Clearly, any of its members are non-zero, and moreover they belong to the larger set C. Hence, C does not only contain the zero vector, and so Y is unbounded.

The case where one assumes that Y is bounded is treated similarly.