Chalmers/Gothenburg University
Mathematical Sciences

EXAM

TMA947/MMG620
OPTIMIZATION, BASIC COURSE

Date:	10–04–06
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:	Adam Wojciechowski (0703-088304)
Result announced:	10–04–23 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:

$$\begin{aligned}
\text{minimize} \quad & z = -2x_1 + x_2 \\
\text{subject to} \quad & x_1 - 3x_2 \leq \beta, \\
& 0 \leq x_1, \\
& 0 \leq x_2 \leq 2.
\end{aligned}$$

(2p) a) Solve this problem for $\beta = -3$ 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 any additional linear programs, state, with a clear motivation, the marginal change in the optimal objective function value when β is varied from its current value of -3 (i.e., state the partial derivative of z^* with respect to β).

Question 2

(implications in theorems)

The following questions consider optimality conditions and theorems related to them. Your task is to find counter-examples showing that some theorems formulated in terms of a given implication is only valid in that direction, and not in the reverse direction.

(1p) a) An elementary result for linear programs (LPs) says that if the program has a finite optimal solution, then there exists an optimal solution among the extreme points. Show that it is not necessary for an optimal solution to an LP to be an extreme point by constructing a specific LP counter-example.[1]

(1p) b) An optimality condition for twice differentiable functions is that if $\nabla f(\boldsymbol{x}^*) =$

[1]The problem formulation in the exam was unclear, and has been modified here.

$\mathbf{0}$ and also $\nabla^2 f(\boldsymbol{x}^*)$ is positive definite, then $\boldsymbol{x}^*$ is a strict local minimum of f on $\mathbb{R}^n$. By presenting a counter-example, show that the reverse implication does not hold true in general

(1p) c) Show, by presenting a specific problem of the form $\min_{x \in S} f(\boldsymbol{x})$, where f is differentiable and S is convex, that the variational inequality is not a sufficient condition for local optimality (it is only a necessary condition). (Recall that the variational inequality considers scalar products of the gradient of f at the point of interest and vectors from the point into the set.)

(3p) Question 3

(modelling)

A cylindrical heat storage unit of diameter D and height H is to be constructed. The heat loss due to convection is $h_c = k_c A(T - T_O)$ and due to radiation is $h_r = k_r A(T^4 - T_O^4)$, where k_c and k_r are constants, A is the surface area of the heat storage, T is the temperature inside the heat storage and T_O is the outside temperature. The heat energy stored in the unit is given by $Q = kV(T - T_O)$, where k is a constant and V is the volume of the heat storage.

Formulate an optimization problem for finding the dimensions of a heat storage such that the heat loss is minimized, at least a given constant Q' of heat is stored, and the storage fits inside a sphere of radius R. Your variables, constants, constraints and objective function should be clearly defined.

Is your model best described as a linear programming, nonlinear programming or mixed integer programming model?

Question 4

(linear programming duality)

Consider the linear programming problem to

$$
\begin{aligned}
\text{minimize} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\
\text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\
& \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u},
\end{aligned}
$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{c}$, $\boldsymbol{l}$, $\boldsymbol{u}$, and $\boldsymbol{x}$ are vectors in $\mathbb{R}^n$, and $\boldsymbol{b} \in \mathbb{R}^m$.

(1p) a) Give the LP dual of this problem.

(1p) b) Prove that this LP dual always has feasible solutions.

(1p) c) What can you conclude if the primal problem has feasible solutions?

Question 5

(Newton's algorithm)

An engineer has decided to verify numerically that the exponential function $x \mapsto \exp(x) = e^x$ grows faster than any polynomial. In order to do so he/she studies the optimization problem to

$$\text{minimize } f(x) = x^\alpha - exp(x), \tag{1}$$

where α is the highest power of the polynomial (we assume it is an even, positive integer number). The engineer uses a Newton method (with unit steps!) to solve the problem. He/she argues that if the exponential function grows faster than any polynomial, then the sequence $\{x_k\}$ generated by the method should converge to infinity, because the objective function f can be decreased indefinitely by increasing the value of x.

(1p) a) State the Newton iteration explicitly for the given problem (1).

(1p) b) Construct a numerical example (that is, choose a value of $\alpha \in \{2, 4, \ldots\}$ and a starting point of the Newton algorithm) illustrating the engineer's error in reasoning.

(1p) c) Find the error in the engineer's reasoning and formally explain it.

(3p) Question 6

(fundamental theorem of global optimality)

Consider the problem to

$$\text{minimize } f(\boldsymbol{x}), \tag{1a}$$

$$\text{subject to } \boldsymbol{x} \in S, \tag{1b}$$

where $S \subseteq \mathbb{R}^n$ is a nonempty set and $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a given function.

Establish the validity of the following theorem.

Consider the problem (1)*, where S is a convex set and f is convex on S. Then, every local minimum of f over S is also a global minimum.*

Question 7

(optimality conditions)

Consider the problem to project (according to the standard Euclidean distance) the vector $\boldsymbol{z} = (2, 3/2)^{\mathrm{T}}$ onto the set S specified by the constraints that $x_j \geq 0$ for $j = 1, 2$, and that $x_1 + x_2 \leq 3/2$.

(1p) a) Describe the appropriate optimization problem to be solved in order to find this projection, and establish that it is a convex problem with a strictly convex objective function.

(1p) b) State the KKT conditions corresponding to a feasible vector $\boldsymbol{x}^*$ being stationary in the problem in a). Establish whether or not the KKT conditions are necessary for a local minimum at $\boldsymbol{x}^*$, and also whether the KKT conditions are sufficient for a feasible vector $\boldsymbol{x}^*$ satisfying the KKT conditions to be a global minimum of the same problem.

(1p) c) Establish whether or not the vector $\boldsymbol{x} = (1, 1/2)^{\mathrm{T}}$ is the projection of $\boldsymbol{z}$ onto the set S.

Good luck!

Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date: 10–04–06

Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) We first rewrite the problem on standard form. In the first constraint we change sign and subtract a non-negative slack (surplus) variable. The upper bound on x_2 is considered as a linear constraint, and in this constraint a second slack variable is added. We get

$$\begin{array}{lrcl} \text{minimize} & z = -2x_1 + x_2, & & \\ \text{subject to} & -x_1 + 3x_2 - s_1 & = & 3, \\ & x_2 + s_2 & = & 2, \\ & x_1, \quad x_2, \quad s_1, \quad s_2 & \geq & 0. \end{array}$$

In phase I, an artifical variable, $a \geq 0$ is added in the first constraint. s_2 is used as the second basic variable. The phase I problem is

$$\begin{array}{lrcl} \text{minimize} & w = a, & & \\ \text{subject to} & -x_1 + 3x_2 - s_1 + a & = & 3, \\ & x_2 + s_2 & = & 2, \\ & x_1, \quad x_2, \quad s_1, \quad s_2, \quad a & \geq & 0, \end{array}$$

and our starting BFS is $(a, s_2)^{\mathrm{T}}$. A calculation of the vector of reduced costs for the non-basic variables x_1, x_2 and s_1 gives $(1, -3, 1)^{\mathrm{T}}$, and hence, x_2 is chosen as the incoming variable. The minimum ratio test shows that a should leave the basis. Since there are no artifical variables left in the basis we have $w = 0$ which is optimal in the phase I problem and which corresponds to a BFS to the original problem. We return to the original problem using $B = (x_2, s_2)^{\mathrm{T}}$, $N = (x_1, s_1)$. The vector of reduced costs are calculated to be $(-5/3, 1/3)^{\mathrm{T}}$ and therefore x_1 is chosen as the incoming variable. The minimum ratio test shows that s_2 should be removed from the basis. Updating B to $(x_2, x_1)^{\mathrm{T}}$ and N to $(s_2, s_1)^{\mathrm{T}}$ and calculating the new reduced costs shows that $\tilde{\boldsymbol{c}}_N = (5, 2) > \boldsymbol{0}$ and hence the current basis is optimal. We have $B^{-1}\boldsymbol{b} = (2, 3)$, i.e., $\boldsymbol{x}^* = (x_1, x_2)^* = (3, 2)$, and $z^* = -4$.

(1p) b) From strong duality we have $z^* = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x^*} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y^*}$ where $\boldsymbol{y}^*$ is the vector of optimal dual variables. The optimal basis will not change for sufficiently small changes of β since $\tilde{\boldsymbol{c}}_N > \boldsymbol{0}$ implies that the optimal basis is unique at the current point. Therefore we have that $\frac{\partial z}{\partial \beta}$ equals the value of the dual variable corresponding to the first con straint. The expression for the dual variables is $\boldsymbol{y}^{\mathrm{T}} = \boldsymbol{c}_B^{\mathrm{T}} B^{-1}$ (which, if one does not remember is clear from the fact that $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^* = \boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{x}_B^* = \boldsymbol{c}_B^{\mathrm{T}} B^{-1}\boldsymbol{b} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y^*}$). We have $\boldsymbol{c}_B^{\mathrm{T}} B^{-1} = (2, -5)^{\mathrm{T}}$, and therefore the marginal change of z^* is twice the change of β.

Question 2

(implications in theorems)

(1p) a) Any LP with $\boldsymbol{c} = \boldsymbol{0}$ would do (as long as the feasible polyhedron has a non-empty relative interior). Here all feasible points are optimal, but not all of them are extreme points. Another simple example is given by $\min x_1$ s.t. $0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1$. Here, $(0, \frac{1}{2})^{\mathrm{T}}$ is an optimal solution, but it is not extreme since, e.g., $(0, \frac{1}{2})^{\mathrm{T}} = \frac{1}{2}(0,0)^{\mathrm{T}} + \frac{1}{2}(0,1)^{\mathrm{T}}$.

(1p) b) Let $n = 1$ and $f(x) = x^4$. Here, $x^* = 0$ is clearly a strict local minimum (and also the global minimum), however the hessian $\nabla^2 f(\boldsymbol{x}^*)$ is not positive definite but only positive semi-definite ($\nabla^2 f(\boldsymbol{x}^*) = 0$).

(1p) c) Let, e.g., $f(x) = -x^2$ and $S = [0, 1]$ for a minimization problem. Here, $x = 0$ fulfills the variational inequality ($\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \geq 0,\ \forall \boldsymbol{x} \in S$), but it is not locally optimal.

Question 3

(modeling)

The model uses the following constants: k_c, k_r, k, T, T_O. The model uses the following variables: D - cylinder diameter, H - cylinder height. The surface area of the cylinder can be expressed as

$$H\pi D + 2\pi \left(\frac{D}{2}\right)^2 = \pi H D + \frac{\pi}{2} D^2.$$

The volume of the cylinder can be expressed as

$$H\pi \left(\frac{D}{2}\right)^2 = \frac{\pi}{2} H D^2.$$

To fit the cylinder inside a sphere, we should place the center of the cylinder at the center of the sphere. The point inside the cylinder most distant from the center is then $\sqrt{(H/2)^2 + (D/2)^2}$ length units away. This must thus be smaller than the radius R. The full model is then to

$$\text{minimize} \quad \left(k_c(T - T_O) + k_r(T^4 - T_O^4)\right)\left(\pi H D + \frac{\pi}{2} D^2\right)$$

$$
\begin{aligned}
\text{subject to} \quad k(T - T_O)\frac{\pi}{2}HD^2 &\geq Q', \\
\sqrt{(H/2)^2 + (D/2)^2} &\leq R, \\
H, D &\geq 0.
\end{aligned}
$$

The model is a nonlinear programming model.

Question 4

(linear programming duality)

(1p) a) The linear programming dual problem is to

$$
\begin{aligned}
\text{maximize} \quad & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\alpha} + \boldsymbol{l}^{\mathrm{T}}\boldsymbol{\beta} - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{\gamma}, \\
\text{subject to} \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\alpha} + \boldsymbol{\beta} - \boldsymbol{\gamma} = \boldsymbol{c}, \\
& \boldsymbol{\beta}, \boldsymbol{\gamma} \geq \boldsymbol{0}^n,
\end{aligned}
$$

where $\boldsymbol{\alpha} \in \mathbb{R}^m$ is the vector of dual variables for the linear constrants, and $\boldsymbol{\beta} \in \mathbb{R}^n$ and $\boldsymbol{\gamma} \in \mathbb{R}^n$ respectively are the vector of dual variables for the lower and upper bounds on $\boldsymbol{x}$.

(1p) b) Set, for example, $\boldsymbol{\alpha} \in \boldsymbol{0}^m$. Then, study the sign of each element of the vector $\boldsymbol{c}$: if $c_j = 0$, set $\beta_j = \gamma_j = 0$; if $c_j > 0$, set $\beta_j = c_j$ and $\gamma_j = 0$; finally, if if $c_j < 0$, set $\beta_j = 0$ and $\gamma_j = -c_j$. This then constitutes a feasible solution to the linear programming dual problem.

(1p) c) The conclusion is that the primal problem has a finite optimal solution; see Theorem 10.6, for example.

Question 5

(Newton's method)

(1p) a) Newton's equation:

$$
x_{k+1} = x_k - \frac{\alpha x^{\alpha-1} - \exp(x)}{\alpha(\alpha-1)x^{\alpha-2} - \exp(x)}.
$$

(1p) b) Probably the simplest counter-example is obtained by taking $x_0 = 1, \alpha = 2$. These initial values cause the Newton's method to generate an oscillating sequence $x_{2k-1} = 0, x_{2k} = 1, k = 1, 2, \ldots$.

(1p) c) The objective function of the problem is not convex in general [may be verified by analyzing the sign of the Hessian $\alpha(\alpha - 1)x^{\alpha-2} - \exp(x)$]. Since the convergence of the Newton method is local in nature, the method is most likely to converge to the nearest local minimum (or maximum if the hessian is negative definite). The engineer thus wrongly assumes the global convergence of the Newton method on non convex functions.

(3p) Question 6

(fundamental theorem of global optimality)

See Theorem 4.3 in The Book.

Question 7

(optimality conditions)

(1p) a) With $\boldsymbol{z} = (2, 3/2)^{\mathrm{T}}$, the appropriate problem to solve is that to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{z}\|^2, \\ \text{subject to} \quad & x_1 + x_2 \le 3/2, \\ & x_j \ge 0, \qquad j = 1, 2. \end{aligned}$$

The objective function is strictly convex: $\nabla f(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{z}$ and $\nabla^2 f(\boldsymbol{x}) = \boldsymbol{I}^n$, where $\boldsymbol{I}^n$ is the identity matrix, so the Hessian $\nabla^2 f(\boldsymbol{x})$ matrix is positive definite everywhere. The problem is a convex one, since also the feasible set is convex—it is indeed a polyhedron.

(1p) b) Changing sign of the second group of constraints, and introducing the Lagrange multiplier vector $\boldsymbol{\mu} \in \mathbb{R}^3$, we obtain the KKT conditions for a feasible vector $\boldsymbol{x}^*$ as follows:

$$\begin{aligned} \boldsymbol{x}^* - \boldsymbol{z} + \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \mu_1(x_1^* + x_2^* - 3/2) &= 0, \\ \mu_2 x_1^* &= 0, \\ \mu_3 x_2^* &= 0. \end{aligned}$$

As the constraints of the problem in a) are affine, the Abadie constraint qualification (CQ) is satisfied; therefore, the KKT conditions are necessary for a local minimum at $\boldsymbol{x}^*$.

As was established in a) above, the optimization problem is a convex one. The above KKT conditions then are sufficient for a feasible vector $\boldsymbol{x}^*$ to be a global minimum of the above problem.

(1p) c) At the given vector $\boldsymbol{x} = (1, 1/2)^{\mathrm{T}}$, it is clear that at any KKT point, $\mu_2 = \mu_3 = 0$ must hold, while complementarity leaves μ_1 free. The remaining linear equation becomes:

$$\mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix};$$

hence, $\mu_1 = 1$. The KKT conditions are satisfied; the vector $\mu^* = (1, 0, 0)^{\mathrm{T}}$ is a vector of Lagrange multipliers, corresponding to the optimal solution $\boldsymbol{x}^* = (1, 1/2)^{\mathrm{T}}$.