Chalmers/Gothenburg University
Mathematical Sciences

EXAM

TMA947/MMG620
OPTIMIZATION, BASIC COURSE

Date:	09–04–14
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 (0762-721860)
Result announced:	09–05–07
	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:

$$
\begin{array}{lrcl}
\text{minimize} & z = 4x_1 + 2x_2 + x_3, & & \\
\text{subject to} & 2x_1 + x_3 & \geq & 3, \\
 & 2x_1 + 2x_2 + x_3 & = & 5, \\
 & x_1, \quad x_2, \quad x_3 & \geq & 0.
\end{array}
$$

(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) Suppose the problem above describes the cost for a production plan, and that the constraints correspond to some sort of demands that have to be satisfied. Motivate (without re-solving the problem!) how much it would be worth to decrease each of the right-hand side components (one by one) with a small number $\varepsilon > 0$.

(3p) Question 2

(optimality conditions)

Farkas' Lemma can be stated as follows:

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{b}$ an $m \times 1$ vector. Then exactly one of the systems

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b}, \\
\boldsymbol{x} &\geq \boldsymbol{0}^n,
\end{aligned} \tag{I}
$$

and

$$
\begin{aligned}
\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} &\leq \boldsymbol{0}^n, \\
\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} &> 0,
\end{aligned} \tag{II}
$$

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

Prove Farkas' Lemma.

(3p) Question 3

(the Frank–Wolfe algorithm)

As applied to the problem of minimizing a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ over a non-empty and bounded polyhedral set $X \subset \mathbb{R}^n$, the Frank–Wolfe method is defined, in short, thus: provide a first feasible solution $\boldsymbol{x}_0$ to the problem, and let $k := 0$; for given $\boldsymbol{x}_k$, solve the LP problem to minimize $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{y}$ over $\boldsymbol{y} \in X$, and let $\boldsymbol{y}_k$ be an optimal solution to this problem. If the value of $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k - \boldsymbol{x}_k)$ is (near) zero, then terminate with $\boldsymbol{x}_k$ being a (near-)stationary point, otherwise let $\boldsymbol{p}_k := \boldsymbol{y}_k - \boldsymbol{x}_k$ and perform a line search in the value of f along the direction $\boldsymbol{p}_k$ from $\boldsymbol{x}_k$, with a maximum step length of 1. Let the resulting vector be $\boldsymbol{x}_{k+1} := \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k$, where α_k is the step length obtained in the line search. Let finally $k := k + 1$, and repeat.

Consider the following constrained nonlinear minimization problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(\boldsymbol{x}) := \frac{1}{2}(x_1^2 + x_2^2), \\ \text{subject to} \quad & \begin{cases} -1 \leq x_1 \leq 2, \\ \ \ \ 1 \leq x_2 \leq 1. \end{cases} \end{aligned}$$

Starting at the point $\boldsymbol{x}_0 = (2, 1)^{\mathrm{T}}$, perform one step of the Frank–Wolfe method. (Feel free to plot the problem and perform the algorithmic steps graphically, as long as you describe all the steps in detail with mathematical notation also.)

Is the point obtained an optimal solution? Why/why not? Explain in detail.

(3p) Question 4

(modeling)

A chocolate producer wants to plan the yearly production of chocolate. He has to fulfill the demand for chocolate in each month according to Table 1. In order to produce 1 kg of chocolate he needs 0.7 kg cocoa and 0.3 kg sugar. He has the possibility to sign a deal with an importer for monthly deliveries of sugar and cocoa; the importer will then deliver the same amount of sugar and cocoa each month (the amount is decided by the chocolate producer, but has to be equal for all months). He can also buy the goods for a higher price at the local market, but has then the possibility to buy different amounts each month. The prices are presented in Table 2. If there are goods left after a month's production, they can be stored until the next month. There is, however, a maximal storage capacity of 100 kg.

Introduce appropriate constants and variables, and create a Linear Programming model that minimizes the yearly production costs.

jan	feb	mar	apr	may	jun	jul	aug	sep	okt	nov	dec
300	230	270	500	150	170	140	230	300	270	350	700

Table 1: Demand of chocolate in kg for each month.

	import	market
cocoa	50	70
sugar	10	12

Table 2: Prices of goods in SEK/kg from import and local market.

(3p) Question 5

(gradient projection)

The gradient projection algorithm is a generalization of the steepest descent method to problems defined over convex sets. Given a point $\boldsymbol{x}_k$ the next point is obtained according to $\boldsymbol{x}_{k+1} = \text{Proj}_X[\boldsymbol{x}_k - \alpha\nabla f(\boldsymbol{x}_k)]$, where X is the convex set over which we minimize, α is the step length and $\text{Proj}_X(\boldsymbol{y}) = \arg\min_{\boldsymbol{x}\in X} \|\boldsymbol{x}-\boldsymbol{y}\|$ (i.e., the closest point to $\boldsymbol{y}$ in X). Note that if $X = \mathbb{R}$ then the method reduces to steepest descent.

Consider the optimization problem to

$$\begin{aligned} &\text{minimize} && f(\boldsymbol{x}) = (x_1 + x_2)^2 + 3(x_1 - x_2)^2, \\ &\text{subject to} && (x_1 - 1)^2 + (x_2 - 2)^2 \leq 1. \end{aligned}$$

Start at the point $x^0 = (1\ 2)^{\mathrm{T}}$ and perform two iterations of the gradient projection algorithm using step length $\alpha = 1/4$. Note that the special form of the feasible region X makes projection very easy! Is the point obtained a global/local optimum? Motivate why/why not!

(3p) Question 6

(a simple optimization problem)

In a recent optimization exam at a Swedish technical university, the following optimization problem was addressed:

$$\begin{aligned} \text{maximize} \quad & f(\boldsymbol{x}) := \sum_{j=1}^{n} a_j/x_j, \\ \text{subject to} \quad & \sum_{j=1}^{n} \log x_j \leq b, \\ & \quad x_j > 0, \quad i = 1, \ldots, n, \end{aligned}$$

where $a_j > 0$ for all j and $b > 0$.

The students were asked to derive the optimal solution to this problem through a Lagrangian relaxation of the first constraint, and by then solving the resulting dual problem. Explain what is wrong with that exam question. In other words, prove that there does not exist an optimal solution to this problem.

[*Hint:* Utilize the KKT conditions.]

(3p) Question 7

(polyhedral theory – LP duality)

Consider the polyhedron $P := \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}^n \,\}$, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Suppose that P is non-empty and that it is contained in some ball, i.e., $\exists M \in \mathbb{R}$ such that $P \subset \{\, \boldsymbol{x} \in \mathbb{R}^n \mid ||\boldsymbol{x}|| \leq M \,\}$.

Show that the polyhedron $Q \subset \mathbb{R}^{n+m}$ defined as all points $\boldsymbol{z} \in \mathbb{R}^{n+m}$ fulfilling the inequalities below is non-empty for all vectors $\boldsymbol{c} \in \mathbb{R}^n$.

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0}^{m\times m} \\ \boldsymbol{0}^{m\times n} & \boldsymbol{A}^{\mathrm{T}} \\ -\boldsymbol{I}^{n\times n} & \boldsymbol{0}^{n\times m} \\ \boldsymbol{0}^{m\times n} & \boldsymbol{I}^{m\times m} \\ \boldsymbol{c}^{\mathrm{T}} & -\boldsymbol{b}^{\mathrm{T}} \end{pmatrix} \boldsymbol{z} \leq \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{c} \\ \boldsymbol{0}^n \\ \boldsymbol{0}^m \\ 0 \end{pmatrix}$$

Good luck!

Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date: 09–04–14
Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) A non-negative slack (surplus) variable is subtracted from the first constraint to transform the problem into standard form.

$$\begin{aligned}
\text{minimize} \quad & z = 4x_1 + 2x_2 + x_3, \\
\text{subject to} \quad & 2x_1 \qquad\quad + x_3 - s_1 = 3, \\
& 2x_1 + 2x_2 + x_3 \qquad = 5, \\
& x_1, \quad x_2, \quad x_3, \quad s_1 \geq 0.
\end{aligned}$$

We start by formulating a phase 1 problem with an artificial variable $a_1 \geq 0$ added in the first constraint. x_2 can be used as a second basic variable.

$$\begin{aligned}
\text{minimize} \quad & w = \qquad\qquad\qquad\qquad a_1, \\
\text{subject to} \quad & 2x_1 \qquad\quad + x_3 - s_1 + a_1 = 3, \\
& 2x_1 + 2x_2 + x_3 \qquad\qquad = 5, \\
& x_1, \quad x_2, \quad x_3, \quad s_1, \quad a_1 \geq 0.
\end{aligned}$$

We start with the BFS given by $(x_2, a_1)^{\mathrm{T}}$. In the first iteration of the simplex algorithm, x_1 has the least reduced cost (-2) and is chosen as the incoming variable. The minimum ration test shows that a_1 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 $\boldsymbol{x}_B = (x_2, x_1)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_3, s_1)^{\mathrm{T}}$ and the reduced costs with the phase 2 cost vector are $\tilde{\boldsymbol{c}}^{\mathrm{T}}_{(x_3,s_1)} = (-\frac{1}{2}, \frac{1}{2})$. The reduced cost for x_3 is negative and x_3 is chosen to enter the basis. $\boldsymbol{B}^{-1}\boldsymbol{b} = (1, \frac{3}{2})^{\mathrm{T}}$ and $\boldsymbol{B}^{-1}\boldsymbol{N}_{x_3} = (0, \frac{1}{2})^{\mathrm{T}}$, therefore x_1 should leave the basis. Updating the basis and computing the new reduced costs gives that $\tilde{\boldsymbol{c}}^{\mathrm{T}}_{(x_1,s_1)} = (2, 0) \geq \boldsymbol{0}$ and thus the optimality condition is fulfilled for the current basis. We have $\boldsymbol{x}^*_B = (1, 3)^{\mathrm{T}}$, or in the original variables, $\boldsymbol{x}^* = (x_1, x_2, x_3)^* = (0, 1, 3)^{\mathrm{T}}$, with the optimal value $z^* = 5$.

(1p) b) The marginal improvement (for a non-degenerate optimal solution) when modifying the right-hand-side vector is given by the values of the dual variables (the "shadow prices"). These are given by $\boldsymbol{y}^* = \boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1} = (0, 1)^{\mathrm{T}}$. Hence, decreasing the first constraint with ϵ gives us nothing. Decreasing the second constraint with ϵ gives an improvement of ϵ of the objective function value.

Question 2

(optimality conditions)

See The Book, Theorem 10.10.

Question 3

(the Frank–Wolfe algorithm)

At $\boldsymbol{x}_0 = (2,1)^{\mathrm{T}}$, $\nabla f(\boldsymbol{x}_0) = (2,1)^{\mathrm{T}}$. Minimizing $2y_1 + y_2$ over the feasible set yields $\boldsymbol{y}_0 = (-1,1)^{\mathrm{T}}$. The search direction therefore is $\boldsymbol{p}_0 = \boldsymbol{y}_0 - \boldsymbol{x}_0 = (-3,0)^{\mathrm{T}}$. The one-dimensional problem (or, line search) then is to minimize $\varphi(\alpha) = f(\boldsymbol{x}_0 + \alpha\boldsymbol{p}_0) = \frac{1}{2}(2-3\alpha)^2 + \frac{1}{2}$, over $\alpha \in [0,1]$. Setting $\varphi'(\alpha) = 0$ yields $\alpha = \frac{2}{3}$; this must be the optimal solution to the line search problem as it belongs to $[0,1]$ and φ is a convex function; the latter holds particularly since f itself is convex. We then obtain, with $\alpha_0 = \frac{2}{3}$, that $\boldsymbol{x}_1 = \boldsymbol{x}_0 + \alpha_0\boldsymbol{p}_0 = (0,1)^{\mathrm{T}}$.

To check whether $\boldsymbol{x}_1$ is optimal, we can, for example, investigate the variational inequality. We have that $\nabla f(\boldsymbol{x}_1) = (0,1)^{\mathrm{T}}$. At $\boldsymbol{x}_1$, all feasible directions are of the form $\{\, \boldsymbol{p} \in \mathbb{R}^2 \mid p_1 \in \mathbb{R},\ p_2 = 0 \,\}$. Hence, for all feasible directions $\boldsymbol{p}$ we have that $\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}\boldsymbol{p} = 0$, and the variational inequality for the problem at hand is fulfilled at $\boldsymbol{x}_1$. Since the problem is convex, $\boldsymbol{x}_1 = (0,1)^{\mathrm{T}}$ must be an optimal solution. [We can also utilize the upper and lower bounds on the optimal value f^* supplied by the algorithm, to reach the same conclusion.]

(3p) Question 4

(modeling)

Introduce the constants d_i for the demand of chocolate in month $i = 1, \ldots, 12$. Let c_1 be the price of 1 kg cocoa from the importer and f_1 the price of 1 kg of cocoa from the market; let c_2 and f_2 be the corresponding prices of sugar. Let a_1 be the amount of cocoa needed for 1 kg of chocolate and a_2 the amount of sugar. Finally let b be the maximal storage capacity.

Introduce the variables x_1 and x_2 for the amount of cocoa/sugar bought from the importer each month. Let y_{1i} and y_{2i} be the amount of cocoa/sugar bought at the local market for months $i = 1, \ldots, 12$. Finally, let z_{1i} and z_{2i} be the amount of cocoa/sugar left in storage after the production in month $i = 1, \dots, 12$ has been completed. The problem is:

$$\min \quad \sum_{i=1}^{12}\sum_{j=1}^{2} c_j x_j + f_j y_{ji},$$

subject to the constraints

$$\begin{aligned}
(x_j + y_{j1} - z_{j1}) &\geq a_j d_1, \quad j = 1, 2\\
(x_j + y_{ji} + z_{ji-1} - z_{ji}) &\geq a_j d_i, \quad i = 2, \dots, 12, j = 1, 2,\\
\sum_{j=1}^{2} z_{ji} &\leq b, \quad i = 1, \dots, 12,\\
x_j, z_{ji}, y_{ji} &\geq 0 \quad i = 1, \dots, 12, \quad j = 1, 2.
\end{aligned}$$

(3p) Question 5

(gradient projection)

Note first that the feasible region X is a circle with center $\boldsymbol{x}_C = (1\ 2)^{\mathrm{T}}$ and radius $r = 1$. Projecting a point $\boldsymbol{y}$ on X results in taking a step of length r in the direction from $\boldsymbol{x}_C$ to $\boldsymbol{y}$. That is:

$$\mathrm{Proj}_X(\boldsymbol{y}) = \begin{cases} \boldsymbol{y} & \text{if } \boldsymbol{y} \in X \\ \boldsymbol{x_C} + r\frac{\boldsymbol{y}-\boldsymbol{x}_C}{\|\boldsymbol{y}-\boldsymbol{x}_c\|} & \text{if } \boldsymbol{y} \notin X \end{cases}$$

The gradient is

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2(x_1 + x_2) + 6(x_1 - x_2) \\ 2(x_1 + x_2) - 6(x_1 - x_2) \end{pmatrix}. \tag{1}$$

Iteration 1: $\boldsymbol{x}^0 = (1\ 2)^T$, $\nabla f(\boldsymbol{x}^0) = (0\ 12)^{\mathrm{T}}$. $\boldsymbol{x}^0 - \alpha\nabla f(\boldsymbol{x}^0) = (1\ 2)^{\mathrm{T}} - (0\ 3)^{\mathrm{T}} = (1\ -1)^{\mathrm{T}}$. $\mathrm{Proj}_X(1\ -1)^{\mathrm{T}} = (1\ 2)^{\mathrm{T}} - (0\ 1)^{\mathrm{T}} = (1\ 1) = \boldsymbol{x}^1$.

Iteration 2: $\boldsymbol{x}^1 = (1\ 1)^T$, $\nabla f(\boldsymbol{x}^1) = (4\ 4)^{\mathrm{T}}$. $\boldsymbol{x}^1 - \alpha\nabla f(\boldsymbol{x}^1) = (1\ 1)^{\mathrm{T}} - (1\ 1)^{\mathrm{T}} = (0\ 0)^{\mathrm{T}}$. $\mathrm{Proj}_X(0\ 0)^{\mathrm{T}} = (1\ 2)^{\mathrm{T}} - \frac{(1\ 2)^{\mathrm{T}}}{\sqrt{5}} = \frac{1}{\sqrt{5}}\binom{\sqrt{5}-1}{2\sqrt{5}-2} = \boldsymbol{x}^2$.

We have convex constraints with an interior point, hence Slaters CQ imply that KKT is necessary for local optimality. The constraint g is active. $\nabla f(\boldsymbol{x}^2) = \binom{0}{12(1-\frac{1}{\sqrt{5}})}$ and $\nabla g(\boldsymbol{x}^2) = \frac{1}{\sqrt{5}}\binom{-2}{-4}$, i.e., they are not parallel. Hence $\boldsymbol{x}^2$ is not a KKT point, and therefore it is not a local (nor a global) minimum.

(3p) Question 6

(a simple optimization problem)

The KKT conditions for this problem amount, apart from complementarity and primal feasibility, to finding a solution in the pair $(\boldsymbol{x}, \mu)^{\mathrm{T}} \in \mathbb{R}^n \times \mathbb{R}_+$ to the nonlinear equations formed by the stationarity conditions for the Lagrangian with respect to $\boldsymbol{x}$, that is, for all $j = 1, \dots, n$,

$$\frac{a_j}{x_j^2} + \frac{\mu}{x_j} = 0.$$

This is clearly impossible, as $x_j > 0$ must be fulfilled, and $a_j > 0$ holds. We therefore conclude that there are not KKT points for this problem.

Can there be optimal solutions that are not KKT points? No, because the linear independence CQ (LICQ) is fulfilled for this problem, so the KKT conditions are necessary conditions for both local and global optimal solutions.

Question 7

(polyhedral theory – LP duality)

Since P is contained in a ball it must be bounded. Also, by assumption it is non-empty. Therefore, for all $\boldsymbol{c} \in \mathbb{R}^n$, there must exist an optimal solution to the problem

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

Then, from the Strong duality theorem it is guaranteed that there is an optimal solution also to the problem

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}, \\ \text{subject to} \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c}, \\ & \boldsymbol{y} \leq \boldsymbol{0}. \end{aligned}$$

The optimal solutions $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ are of course feasible, and we know that $c^{\mathrm{T}}\boldsymbol{x}^* = b^{\mathrm{T}}\boldsymbol{y}^*$ holds (then also $c^{\mathrm{T}}\boldsymbol{x}^* \leq b^{\mathrm{T}}\boldsymbol{y}^*$ holds, which is not true for a general feasible pair). So, with $\boldsymbol{z} = \left((\boldsymbol{x}^*)^{\mathrm{T}}, (\boldsymbol{y}^*)^{\mathrm{T}}\right)^{\mathrm{T}}$ all of the inequalities are fulfilled, and the polyhedron Q is proved to be non-empty.