Chalmers/Gothenburg University
Mathematical Sciences

**EXAM**

# TMA947/MMG620
# OPTIMIZATION, BASIC COURSE

| | |
|---|---|
| **Date:** | 10–12–13 |
| **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.<br>Questions are *not* numbered by difficulty.<br>To pass requires 10 points and three passed questions. |
| **Examiner:** | Michael Patriksson |
| **Teacher on duty:** | Emil Gustavsson (0703-088304) |
| **Result announced:** | 11–01–05<br>Short answers are also given at the end of<br>the exam on the notice board for optimization<br>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:

$$\begin{aligned} \text{maximize} \quad & 3x_1 + x_2, \\ \text{subject to} \quad & 3x_1 + 2x_2 \geq 1, \\ & 2x_1 + x_2 \leq 2, \\ & x_1 \geq 0, \\ & x_2 \in \mathbb{R}. \end{aligned}$$

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

Aid: Utilize the identity

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

**(1p)** b) Is the solution obtained unique? Motivate your answer!

---

## Question 2

(optimality conditions)

Suppose that for $j = 1, \ldots, n$ the functions $f_j : \mathbb{R} \to \mathbb{R}$ are convex and differentiable. Let $b > 0$. Our problem, called the *resource allocation problem*, has the following general statement:

$$\underset{x}{\text{minimize}} \quad \sum_{j=1}^{n} f_j(x_j), \tag{1a}$$

$$\text{subject to} \quad \sum_{j=1}^{n} x_j = b, \tag{1b}$$

$$x_j \geq 0, \qquad j = 1, \ldots, n. \tag{1c}$$

While this problem is exceptionally simple it has many applications, for example in portfolio optimization, production economics, and stratified sampling.

**(2p)** a) Introduce any necessary multipliers, and describe the necessary Karush–Kuhn–Tucker conditions for a vector $\boldsymbol{x}^*$ to be a local optimum in the problem (1). Are these conditions also sufficient for the global optimality of a KKT-point $\boldsymbol{x}^*$?

**(1p)** b) Based on the result in a), establish the following result on the characterization of optimal solutions to the problem (1), known as Gibbs' Lemma: *Suppose that $\boldsymbol{x}^*$ solves the problem* (1)*. Then, there exists (at least one) $\lambda^* \in \mathbb{R}$ such that*

$$f_j'(x_j^*) \begin{cases} = \lambda^*, & \text{if } x_j^* > 0, \\ \geq \lambda^*, & \text{if } x_j^* = 0, \end{cases} \qquad j = 1, \ldots, n, \tag{2}$$

*holds.*

---

## Question 3

(modeling)

A group of $n$ people has decided to arrange a party for New Year's Eve. Each member of the group has purchased things to the party for $d_i$ SEK, $i \in \{1, \ldots, n\}$. Your assignment is to decide how money should be transferred between the members such that all members will have paid equally much.

**(1p)** a) Introduce the necessary variables and formulate a linear programming (LP) model which minimizes the total amount of transfered cash between the members.

**(2p)** b) In the solution to the first model, one of the participants is supposed to give money to six other members; this is fairly impractical. Introduce additional integer variables and extend your model to a linear mixed integer programming model (i.e., the model should be linear if the integrality constraints are relaxed) such that *each member only needs to give money to at most one other member.*

---

**(3p)** **Question 4**

(the Frank–Wolfe method)

Consider the problem to

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

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 fist-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.)

Give the upper and lower bounds of the optimal objective function value that the algorithm generates in each iteration, and give a theoretical motivation for them. If an optimum is found, motivate why it is an optimum.

---

**Question 5**

(Lagrangian duality)

Consider the problem to find the Euclidean projection of the origin in $\mathbb{R}^2$ on the polyhedral set defined by the three linear inequalities $x_1 \leq 4$, $x_2 \leq 4$, and $x_1 + x_2 \geq 4$.

**(1p)** a) State this projection problem as a convex quadratic optimization problem.

**(1p)** b) By Lagrangian relaxing the constraint that $x_1 + x_2 \geq 4$ must hold, formulate the corresponding Lagrangian dual problem *explicitly*. Establish that the Lagrangian dual problem is that of maximizing a concave function.

**(1p)** c) Solve this Lagrangian dual problem. Utilize the primal–dual relationships between the primal and the dual problem to establish the solution to the original problem. Confirm your answer graphically.

---

**(3p)** **Question 6**

(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.

---

**Question 7**

(short questions)

Answer the following three short questions. You *must* motivate your answers in order to receive any points.

**(1p)** a) Consider the following problem

$$\underset{x \in X}{\text{minimize}}\, f(\boldsymbol{x}),$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and $X \subset \mathbb{R}^n$ is a convex and bounded set. Assume further that for some $M \in \mathbb{R}$ $f(\boldsymbol{x}) \geq M$ for all $\boldsymbol{x} \in X$ holds. Does an optimal solution always exist to this problem? If not, give a counter-example!

**(1p)** b) Assume that the objective function $f : \mathbb{R}^2 \to \mathbb{R}$ is a convex function. Define the constraint functions as

$$g_1(x, y) := \begin{cases} (x+1)^2 + (y+1)^2 - 1, & \text{for } x < -1, \\ (y+1)^2 - 1, & \text{for } -1 \leq x \leq 1, \\ (x-1)^2 + (y+1)^2 - 1, & \text{for } 1 < x, \end{cases}$$

and

$$g_2(x, y) := -y.$$

Consider the following problem:

$$\begin{aligned} \text{minimize} \quad & f(x), \\ \text{subject to} \quad & g_1(x, y) \leq 0, \\ & g_2(x, y) \leq 0. \end{aligned}$$

Answer the following two questions about the problem described above. If a feasible point $x$ satisifies the KKT conditions, does it then imply that the point is optimal? If a feasible point $x$ is optimal, does it then imply that it satisfies the KKT conditions?

**(1p)** c) Consider the following set

$$X = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} (x_i - a_i)^2 \leq b, \ \exp\left( \sqrt{\sum_{i=1}^{n} (x_i - c_i)^2} \right) \geq d \right\}.$$

Assume that for the constants $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{c} \in \mathbb{R}^n$, $0 < d \in \mathbb{R}$ and $0 < b \in \mathbb{R}$ the following inequality holds

$$\sqrt{\sum_{i=1}^{n} (a_i - c_i)^2} \geq \sqrt{b} + \ln d.$$

Is the set X convex?

---

*Good luck!*

Chalmers/Gothenburg University
Mathematical Sciences

**EXAM SOLUTION**

# TMA947/MAN280
# OPTIMIZATION, BASIC COURSE

**Date:** 10–12–13

**Examiner:** Michael Patriksson

## Question 1

(the simplex method)

**(2p)** a) We first rewrite the problem on standard form. We multiply the objective by $-1$ to obtain a minimization problem and introduce the variables $x_2^+$ and $x_2^-$ such that $x_2 = x_2^+ - x_2^-$, and slack variables $s_1$ and $s_2$.

$$
\begin{array}{llrrrrrl}
\text{minimize} & z = & -3x_1 & -x_2^+ & +x_2^- & & & \\
\text{subject to} & & 3x_1 & +2x_2^+ & -2x_2^- & -s_1 & & = 1 \\
& & 2x_1 & +x_2^+ & -x_2^- & & +s_2 & = 2 \\
& & x_1, & x_2^+, & x_2^-, & s_1, & s_2 & \geq 0.
\end{array}
$$

In phase I the artificial variable $a$ is added in the first constraint, $s_2$ is used as the second basic variable. We obtain the problem

$$
\begin{array}{llrrrrrrl}
\text{minimize} & w = & & & & & & a & \\
\text{subject to} & & 3x_1 & +2x_2^+ & -2x_2^- & -s_1 & & +a & = 1 \\
& & 2x_1 & +x_2^+ & -x_2^- & & +s_2 & & = 2 \\
& & x_1, & x_2^+, & x_2^-, & s_1, & s_2 & a & \geq 0.
\end{array}
$$

The starting BFS is thus $(a, s_2)^{\mathrm{T}}$. Calculating the vector of reduced costs for the non-basic variables $x_1, x_2^+, x_2^-$ and $s_1$ yields $(-3, -2, 2, 1)^{\mathrm{T}}$. Thus $x_1$ enters the basis. The minimum ratio test shows that $a$ should leave the basis. We thus have a BFS without artificial variables, and may proceed with pha se II.

We have the basic variables $(x_1, s_2)$. The vector of reduced costs for the non-basic variables $x_2^+, x_2^-$ and $s_1$ is $(1, -1, -1)$. We may choose either $x_2^-$ or $s_1$ to enter the basis. We take $x_2^-$. The minimum ratio test implies that $s_2$ must leave the basis. We now have $x_1, x_2^-$ as basic variables. The vector of reduced costs for the non-basic variables $x_2^+, s_1, s_2$ is $(0, 1, 3)^{\mathrm{T}}$. The current point is optimal. We thus have $(x_1, x_2^-, x_2^+, s_1, s_2) = (3, 4, 0, 0, 0)$, or in the original variables, $(x_1, x_2) = (3, -4)$.

**(1p)** b) The reduced costs are not strictly positive; we can thus not conclude that there is a unique optimal solution. We may introduce $x_2^+$ into the basis; the minimum ratio test can however not provide a variable that leaves the basis (all entries are negative in $B^{-1}N_j$). This is because we may let $x_2^+ = \alpha$, $x_2^- = 4 + \alpha$ for all $\alpha \geq 0$ and obtain an optimal solution in the problem written on standard form. All these solutions however correspond to the same solution $(x_1, x_2) = (3, -4)$ in the original problem. The solution in the original problem is unique (which can also be realized by checking that it is the only KKT point).

---

## Question 2

(optimality conditions)

**(2p)** a) Thanks to the linearity of the constraints, the problem satisfies the Abadie constraint qualification and the Karush–Kuhn–Tucker conditions are necessary for the local optimality of $\boldsymbol{x}^*$. As the problem is convex the KKT conditions are also sufficient for $\boldsymbol{x}^*$ to be a global optimum.

Introducing the multiplier $\lambda$ for the equality constraint and $\mu_j \geq 0$ for the sign condition on $x_j$, we obtain the Lagrange function $L(\boldsymbol{x}, \mu, \boldsymbol{\lambda}) := -b\lambda + \sum_{j=1}^{n}(f_j(x_j) - [\lambda + \mu_j]x_j)$. Setting the partial derivatives of $L$ with respect to each $x_j$ to zero yields

$$f'(x_j^*) = \lambda^* + \mu_j^*, \qquad j = 1, \dots, n. \tag{1}$$

Further, the complementarity conditions state that

$$\mu_j^* \cdot x_j^* = 0, \qquad j = 1, \dots, n.$$

Together with the dual feasibility conditions that $\mu_j^* \geq 0$ for all $j$ and that $\boldsymbol{x}^*$ fulfills the primal feasibility conditions that $\boldsymbol{x}^* \geq \boldsymbol{0}^n$ and $\sum_{j=1}^{n} x_j^* = b$, we have stated all the KKT conditions.

**(1p)** b) Suppose that the triple $(\boldsymbol{x}^*, \mu^*, \boldsymbol{\lambda}^*) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$ is a Karush–Kuhn–Tucker point. For a $j$ with $x_j^* > 0$ we must therefore have from (1) that $f'(x_j^*) = \lambda^*$. Suppose instead that $x_j^* = 0$. Then, since $\mu_j^* \geq 0$ must hold, we obtain from (1) that $f'(x_j^*) = \lambda^* + \mu_j^* \geq \lambda^*$, and we are done.

---

## Question 3

(modeling)

**(1p)** a) Introduce the variable $x_{ij}$ for the amount of money person $i$ gives to person

$j$. The model is to

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}, \\
\text{subject to} \quad & d_i + \sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n} x_{ji} = \frac{1}{n}\sum_{j=1}^{n} d_j, \quad i = 1,\dots,n. \\
& x_{ij} \geq 0 \qquad i = 1,\dots,n, \; j = 1,\dots,n.
\end{aligned}$$

**(2p)** b) Introduce the variables $y_{ij}$, where

$$y_{ij} = \begin{cases} 1 & \text{if person } i \text{ gives any money to person } j \\ 0 & \text{otherwise.} \end{cases}$$

Then the model is to

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}, \\
\text{subject to} \quad & d_i + \sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n} x_{ji} = \frac{1}{n}\sum_{j=1}^{n} d_j, \quad i = 1,\dots,n, \\
& \sum_{j=1}^{n} y_{ij} = 1, \qquad i = 1,\dots,n, \\
& x_{ij} \leq M y_{ij}, \qquad i = 1,\dots,n, \; j = 1,\dots,n, \\
& x_{ij} \geq 0, \qquad i = 1,\dots,n, \; j = 1,\dots,n, \\
& y_{ij} \in \{0,1\}, \qquad i = 1,\dots,n, \; j = 1,\dots,n,
\end{aligned}$$

where $M$ is some large number. $M = \sum_{i=1}^{n} d_i$ is large enough.

---

## (3p) Question 4

(the Frank-Wolfe method)

*Iteration 1:* $\boldsymbol{x}_0 = (0,0)^{\mathrm{T}}$ is feasible and $f(\boldsymbol{x}_0) = 0$, so we get: $[LBD, UBD] = (-\infty, 0]$. $\nabla f(\boldsymbol{x}_0) = (-3,-6)^{\mathrm{T}}$ and the solution to the LP $\min_{\boldsymbol{y}} \nabla f(\boldsymbol{x}_0)^{\mathrm{T}}\boldsymbol{y}$ is obtained at $\boldsymbol{y}_0 = (2,2)^{\mathrm{T}}$. 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 $\boldsymbol{y} \in \mathbb{R}^{\not\vDash}$. The LP problem is a relaxation of the original problem, hence an optimal objective value gives a lower bound. The optimal objective value of the LP is $f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^{\mathrm{T}}(\boldsymbol{y}_0 - \boldsymbol{x}_0) = 0 + (-3,-6)^{\mathrm{T}}(2,2) = -18$. Hence, $[LBD, UBD] = [-18, 0]$. The search direction is $\boldsymbol{p}_0 = \boldsymbol{y}_0 - \boldsymbol{x}_0 = (2,2)^{\mathrm{T}}$. Line search: $\phi(\alpha) := f(\boldsymbol{x}_0 + \alpha\boldsymbol{p}_0) = f((2\alpha, 2\alpha)^{\mathrm{T}}) = 12\alpha^2 - 18\alpha$. $\phi'(\alpha) = 24\alpha - 18 = 0 \Rightarrow \alpha = 3/4 < 1$. Hence, $\boldsymbol{x}_1 = (3/2, 3/2)^{\mathrm{T}}$.

*Iteration 2:* $f(\boldsymbol{x}_1) = -27/4$, so $[LBD, UBD] = [-18, -27/4]$. $\nabla f(\boldsymbol{x}_1) = (3/2, -3/2)^{\mathrm{T}}$ and the solution to the LP $\min_{\boldsymbol{y}} \nabla f(\boldsymbol{x}_1)^{\mathrm{T}}\boldsymbol{y}$ is obtained at $\boldsymbol{y}_1 =$

$(1,2)^{\mathrm{T}}$. $f(\boldsymbol{x}_1) + \nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{y}_1 - \boldsymbol{x}_1) = -33/4$, so $[LBD, UBD] = [-33/4, -27/4]$. The search direction is $\boldsymbol{p}_1 = \boldsymbol{y}_1 - \boldsymbol{x}_1 = (-1/2, 1/2)^{\mathrm{T}}$. Line search, $\phi(\alpha) := f(\boldsymbol{x}_1 + \alpha \boldsymbol{p}_1) = f((3/2 - \alpha/2, 3/2 + \alpha/2)^{\mathrm{T}}) = \alpha^2/4 - (6/4)\alpha - 27/4$. $\phi'(\alpha) = 2\alpha/4 - 3/4 = 0 \Rightarrow \alpha = 3 > 1$. Hence, take $\alpha = 1$ and $\boldsymbol{x}_2 = (1,2)^{\mathrm{T}}$.

$\boldsymbol{x}_2 = (1,2)^{\mathrm{T}}$ is a KKT point. The objective function is convex (all eigenvalues to the Hessian are non-negative) and the feasible set is a polyhedron, so the problem is convex. The KKT conditions are sufficient for optimality for convex problems, so $\boldsymbol{x}_2 = (1,2)^{\mathrm{T}}$ is an optimal solution with $f(\boldsymbol{x}_2) = -8$.

---

## Question 5

(Lagrangian duality)

**(1p)** a) The problem can be stated as that to minimize $f(\boldsymbol{x}) := \frac{1}{2}(x_1^2 + x_2^2)$ subject to the constraints that $x_1 + x_2 \geq 4$ and $x_j \leq 4$, $j = 1, 2$.

**(1p)** b) Introducing the Lagrange multiplier $\mu \geq 0$ for the constraint $x_1 + x_2 \geq 4$, the Lagrangian subproblem has the form

$$\underset{x_j \leq 4,\, j=1,2}{\text{minimize}} \; 4\mu + \frac{1}{2}x_1^2 - \mu x_1 + \frac{1}{2}x_2^2 - \mu x_2.$$

The problem separates over each variable, and the solutions are symmetric: for $0 \leq \mu \leq 4$, $x_j = \mu$ for $j = 1, 2$, while for $\mu > 4$, $x_j = 4$ for $j = 1, 2$. The explicit Lagrangian dual function hence is to maximize the function $q$ over $\mu \geq 0$, where $q(\mu) = 4\mu - \mu^2$ for $0 \leq \mu \leq 4$, and $q(\mu) = 16 - 4\mu$ for $\mu \geq 4$. Its derivative hence is $q'(\mu) = 4 - 2\mu$ for $0 \leq \mu \leq 4$, and $q'(\mu) = -4$ for $\mu \geq 4$. The Lagrangian dual function clearly is concave over $\mu \geq 0$.

**(1p)** c) The solution to the Lagrangian dual problem is $\mu^* = 2$. Utilizing the result in b) we may derive that $\boldsymbol{x}^* = (2,2)^{\mathrm{T}}$. Strong duality holds, that is, $f(\boldsymbol{x}^*) = q(\mu^*)$.

---

## (3p) Question 6

(optimality conditions)

See Theorem 10.10.

## Question 7

(short questions)

**(1p)** a) X can be defined as an open set! Define $f(x) = x$ and $X = \{0 < x < 1\}$, the problem does not have an optimal solution.

**(1p)** b) The feasible set is convex (it is the line segment between (-1,0) and (1,0)). Th us KKT is sufficient (first question: yes). The set does not have an interior point, thus slater does not hold. LICQ does no t hold either. The objective $f(x, y) := x + y$ would result in an optimal solution at (-1,0), wh ich is not a KKT point, hence KKT is not necessary (second question: no).

**(1p)** c) We will use the notation $||a|| = \sqrt{\sum_{i=1}^{n} a_i^2}$. Assume that $||x - a||^2 \leq b$. We have that $||a - c|| = ||a - x + x - c|| \leq ||a - x|| + ||x - c||$, where the last inequality is the triangle inequality. Hence $||x - c|| \geq ||a - c|| - ||a - x|| \geq q \ln d + \sqrt{b} - \sqrt{b} = \ln d$. Therefore $\exp(||x - c||) \geq d$. This means that if we satisfy the first constr aint, then the other constraint is automatically satisfied (hence it is redundant). Since the first c onstraint is a convex function, the set is convex.