Chalmers/Gothenburg University
Mathematical Sciences

EXAM

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date:	12–04–10
Time:	House V, morning, 8^{30}–13^{30}
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:	Hossein Raufi (0703-088304)
Result announced:	12–04–25 Short answers are also given at the end of the exam on the notice board for optimization in the MV building.

Exam instructions

<u>When you answer the questions</u>

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.

<u>At the end of the exam</u>

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:

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

(2p) a) Solve this problem using phase I (so that you begin with a unit matrix as the first basis) and phase II of the simplex method.

Aid: Utilize the identity

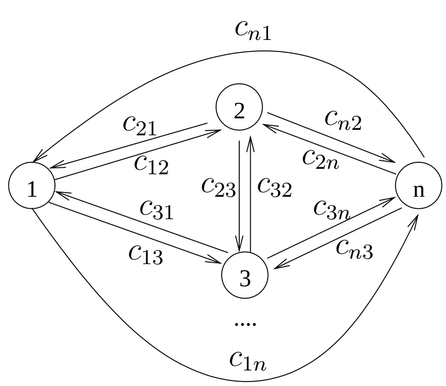
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(1p) b) If an optimal solution exists, use your calculations to decide if it is unique. If the problem is unbounded, use your calculations to specify a direction of unboundedness of the objective value.

(3p) Question 2

(LP duality)

Consider the following complete graph on n nodes:



Let the set $A = \{\, (i,j) \in \{1,\dots,n\} \times \{1,\dots,n\} \mid i \neq j \,\}$ be the set of all arcs in the above graph. Assume given constants $c_{ij} \geq 0$ for each arc $(i,j) \in A$. Introduce variables x_{ij} for all arcs $(i,j) \in A$ and an additional variable s, and consider the following LP problem:

$$
\begin{aligned}
\text{maximize} \quad & s, \\
\text{subject to} \quad & \sum_{\substack{i=1 \\ i\neq j}}^{n} x_{ij} - \sum_{\substack{i=1 \\ i\neq j}}^{n} x_{ji} = \begin{cases} -s, & j = 1, \\ 0, & j \in \{2,\dots,n-1\}, \\ s, & j = n, \end{cases} \\
& 0 \leq x_{ij} \leq c_{ij}, \qquad (i,j) \in A.
\end{aligned} \tag{1}
$$

i) Interpret the LP (1) in terms of the graph, what does it mean?

ii) Find the LP dual to (1).

iii) Interpret the LP dual problem in terms of the graph.

Hint: The concept of a *cut* is important when interpreting the dual. A cut between node $i \in \{1,\dots,n\}$ and node $j \in \{1,\dots,n\}$ corresponds to a set of arcs, such that by removing these, no path between i and j exists in the graph (the graph is cut into two parts).

Question 3

(modeling)

Consider K convex polyhedra $P^1, P^2, \dots, P^K$ in $\mathbb{R}^n$ described by

$$P^k = \{\, \boldsymbol{x} \in \mathbb{R}^n \mid A^k \boldsymbol{x} \leq \boldsymbol{b}^k \,\}, \quad \text{where } A^k \in \mathbb{R}^{m_k \times n},\ \boldsymbol{b}^k \in \mathbb{R}^{m_k} \quad \text{for } k = 1,\dots,K.$$

The *convex hull* of a set P is defined as the smallest convex set containing P and is denoted by $\operatorname{conv}(P)$.

(1p) a) Formulate a linear optimization model for minimizing the objective function $\boldsymbol{c}^T\boldsymbol{x}$ over the set

$$\operatorname{conv}\Big(\bigcap_{k=1}^{K} P^k\Big).$$

(2p) b) Formulate an optimization model for minimizing the objective function $\boldsymbol{c}^T\boldsymbol{x}$

over the set

$$\operatorname{conv}\Big(\bigcup_{k=1}^{K} P^k\Big).$$

The model should only contain continuous variables and continuous constraints. It does however not need to be linear.

Question 4

(exterior penalty method)

Consider the problem to

$$\begin{aligned} &\text{minimize} \quad f(\boldsymbol{x}) := x_1^2 + x_2^2, \\ &\text{subject to} \quad h(\boldsymbol{x}) := x_1 + x_2 - 1 = 0. \end{aligned} \tag{1}$$

We consider solving (1) by using the external penalty method with the quadratic penalty function $\psi(s) := s^2$, $s \in \mathbb{R}$. The penalty problem is to

$$\underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} \quad f(\boldsymbol{x}) + \nu \hat{\chi}_S(\boldsymbol{x}),$$

where $\hat{\chi}_S(\boldsymbol{x}) = \psi(h(\boldsymbol{x}))$, for positive, increasing values of the penalty parameter ν.

(1p) a) By applying the KKT conditions to problem (1), determine its optimal solution and the corresponding optimal Lagrange multiplier.

(1p) b) Apply the exterior penalty method for the problem (1), and show that the sequence of (explicitly stated) solutions to the penalty problem converges to the unique primal solution when $\nu \to \infty$.

(1p) c) Provide the corresponding sequence of estimates of the Lagrange multiplier, and show that it converges to the optimal Lagrange multiplier.

Question 5

(linear programming: existence of optimal solutions)

Let $P := \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{x} \geq \boldsymbol{0}^n \,\}$ and $V := \{\boldsymbol{v}^1, \dots, \boldsymbol{v}^k\}$ be the set of extreme points of P. Further, let $C := \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}^m;\ \boldsymbol{x} \geq \boldsymbol{0}^n \,\}$ and $D := \{\boldsymbol{d}^1, \dots, \boldsymbol{d}^r\}$ be the set of extreme directions of C.

Consider the linear program

$$\begin{aligned} \text{minimize} \quad & z = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} \quad & \boldsymbol{x} \in P. \end{aligned} \tag{1}$$

The task is to establish the following two basic facts about solutions to the basic linear program (1).

(2p) a) This problem has a finite optimal solution if and only if P is nonempty and z is lower bounded on P, that is, if P is nonempty and $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}^j \geq 0$ for all $\boldsymbol{d}^j \in D$.

(1p) b) If the problem has a finite optimal solution, then there exists an optimal solution among the extreme points.

Question 6

(basic facts in optimization)

The below three statements should be assessed. Are they true or false? Provide an answer together with a short but complete motivation.

(1p) a) For the phase I problem of the simplex method (when a BFS is *not* known a priori), the optimal value is always zero.

(1p) b) Suppose $f \in C^2$. If, at some iteration point $\boldsymbol{x} \in \mathbb{R}^n$ there exists a solution $\boldsymbol{p}$ to the search direction-finding problem of Newton's method then it defines a descent direction for f at $\boldsymbol{x}$.

(1p) c) Consider the convex program

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}), \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

where the functions f and g_i, $i = 1, \dots, m$, are convex. Suppose that $\boldsymbol{x}^*$ is a globally optimal solution to this problem, and that $g_k(\boldsymbol{x}^*) < 0$ for some index $k \in \{1, \dots, m\}$. Then, if we remove constraint k from the problem its set of optimal solutions is unchanged.

Question 7

(Lagrangian duality)

Consider the strictly convex quadratic optimization problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := \frac{1}{2}(x_1 - 5)^2 + \frac{1}{2}(x_2 - 3)^2, && \text{(1a)} \\ \text{subject to} \quad & x_1 + x_2 \leq 5, && \text{(1b)} \\ & 0 \leq x_j \leq 3, \quad j = 1, 2. && \text{(1c)} \end{aligned}$$

(1p) a) Establish that the optimal solution to this problem is $\boldsymbol{x}^* = (3, 2)^{\mathrm{T}}$, by utilizing the Karush–Kuhn–Tucker conditions.

(2p) b) Do the following:

[1] Explicitly state its Lagrangian dual function q and its Lagrangian dual problem, associated with the Lagrangian relaxation of the constraint (1b).

[2] Solve this Lagrangian dual problem and provide the optimal Lagrange multiplier μ^*. Confirm that this Lagrange multiplier equals the KKT multiplier corresponding to the inequality (1b), from a).

[3] Prove that strong duality holds, that is, prove that $q(\mu^*) = f(\boldsymbol{x}^*)$ holds.

Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date: 12–04–10

Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) We first rewrite the problem on standard form. We introduce variables x_1^+ and x_1^- and let $x_1 = x_1^+ - x_1^-$. We also add slack variables s_1 and s_2.

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

In phase I the artificial variable a is added in the first constraint, s_2 is used as the second basic variable in order to obtain a unit matrix as the first basis. We obtain the phase I problem

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

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

We have the basic variables (x_2, s_2). The vector of reduced costs for the non-basic variables x_1^+, x_1^- and s_1 is $(3/2, -3/2, -1/2)$. We let x_1^- enter the basis. The minimum ratio test implies that x_2 leaves the basis. We now have x_1^-, s_2 as basic variables. The vector of reduced costs for the non-basic variables x_1^+, x_2 and s_1 is $(0, 3, -2)^{\mathrm{T}}$. Thus s_1 enters the basis. Minimum ratio implies that s_2 leaves basis. We now have x_1^-, s_1 as basic variables. The vector of reduced costs for the non-basic variables x_1^+, x_2 and s_2 is $(0, 1, 2)^{\mathrm{T}}$.

Since the reduced costs are all non-negative, the current BFS is optimal. Returning to the original variables, we obtain $(x_1, x_2) = (-3, 0)$ as the optimal solution and -6 as the optimal value.

(1p) b) In the optimal BFS, the reduced cost corresponding to x_1^+ is zero. Therefore, we can let x_1^+ enter the basis without changing the objective. We do not obtain any leaving variable as minimum ratio implies that the problem is unbounded in that direction. This is simply the increasing x_1^+ and increasing x_1^- by the same amount (which can be any positive number). So the problem in standard form does not have a unique optimal solution, but the problem formulated in the original variables does since all these solutions correspond to $(-3, 0)$. Replacing one free variable with two positive variables always implies that each solution is non-unique in the sense described above.

(3p) Question 2

(LP duality)

i) We interpret the variables x_{ij} as the flow from node i to node j of the graph. The first set of constraints guarantee that the flow going into a node equals the flow leaving that node (i.e. flow balance) except at node 1 where s units of flow enters and at node n where s unit of flow leaves. The second set of constraints imply a limitof flow on all arcs and that flow can not be negative. To summerize, we push s units of flow into the graph at node 1 and take out s units of flow at node n and we maximize s. This can be interpreted as maximizing the flow through the graph from node 1 to node n.

ii) Let v_i for $i \in \{1, \ldots, n\}$ denote the dual variables corresponding to the first set of constraints and w_{ij} for $(i, j) \in A$ denote the dual variables corresponding to the second set of constraints.. The dual then becomes:

$$\begin{aligned}
\text{maximize} \quad & \sum_{(i,j)\in A} w_{ij} c_{ij} \\
\text{subject to} \quad & w_{ij} - v_i + v_j \geq 0, \; (i, j) \in A, \\
& v_1 - v_n = 1, \\
& w_{ij} \geq 0, \; (i, j) \in A.
\end{aligned}$$

iii) We are maximizing with positive costs on w_{ij}, therefore we would want to put them all to zero. Unfortunately, this is not feasible. First, the constraints imply that if $w_{ij} = 0$ then $v_i \leq v_j$. Second $v_1 = v_n + 1$. But by following a path from node 1 to node n throught the graph we obtain that $v_1 \leq v_n$. Therefore, all paths that connect 1 with n need to contain an arc (i, j) where $w_{ij} = 1$ in order to break the chain of inequalities such that $v_1 = v_n + 1$ is possible. This will generate cost c_{ij}. If we let c_{ij} correspond to the cost of including the arc (i, j) in a cut, the dual is to find the minimal cost cut between nodes 1 and n.

Question 3

(modeling)

(1p) a) Since P^k is convex for $k = 1, \dots, K$, we have that

$$\operatorname{conv}\left(\bigcap_{k=1}^{K} P^k\right) = \bigcap_{k=1}^{K} P^k = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, A^k \boldsymbol{x} \leq \boldsymbol{b}^k,\ k = 1, \dots, K \right\}.$$

Hence, we can write the optimization problem as that to

$$\begin{aligned} &\text{minimize} && \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}, \\ &\text{subject to} && A^k \boldsymbol{x} \leq \boldsymbol{b}^k, \quad k = 1, \dots, K. \end{aligned}$$

(2p) b) Any point in $\operatorname{conv}\left(\bigcup_{k=1}^{K} P^k\right)$ can be represented as a convex combination of points in the individual polyhedra $P^1, \dots, P^K$, i.e.,

$$\boldsymbol{x} = \sum_{k=1}^{K} \lambda_k \boldsymbol{x}^k, \quad \text{with} \quad \sum_{k=1}^{K} \lambda_k = 1, \quad \lambda_k \geq 0,$$

where $\boldsymbol{x}^k \in P^k$, $k = 1, \dots, K$. Hence, we can formulate the optimization problem as that to

$$\begin{aligned} &\text{minimize} && \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}, \\ &\text{subject to} && \boldsymbol{x} = \sum_{k=1}^{K} \lambda_k \boldsymbol{x}^k, \\ &&& A^k \boldsymbol{x}^k \leq \boldsymbol{b}^k, \quad k = 1, \dots, K, \\ &&& \sum_{k=1}^{K} \lambda_k = 1, \\ &&& \lambda_k \geq 0, \quad k = 1, \dots, K, \end{aligned}$$

where $\boldsymbol{x}, \boldsymbol{x}^1, \dots, \boldsymbol{x}^K$ and $\lambda_1, \dots, \lambda_K$ are the variables in the optimization model.

(This model can actually be extended to a linear model by the variable substitution $\bar{\boldsymbol{x}}^k = \lambda_k \boldsymbol{x}^k$.)

Question 4

(exterior-penalty methods)

(1p) a) By applying the KKT conditions on the problem, we obtain the unique solution $\boldsymbol{x}^* = (1/2, 1/2)^{\mathrm{T}}$ and $\lambda^* = -1$.

(1p) b) Applying the exterior quadratic penalty method, we get the unconstrained problem

$$\min_{\boldsymbol{x} \in \mathbb{R}^2} \Big(f(\boldsymbol{x}) + \nu h(\boldsymbol{x})^2 \Big) = \min_{\boldsymbol{x} \in \mathbb{R}^2} \Big(x_1^2 + x_2^2 + \nu(x_1 + x_2 - 1)^2 \Big).$$

Setting the gradient to zero we obtain

$$2x_1 + 2\nu x_1 + 2\nu x_2 - 2\nu = 0,$$
$$2x_2 + 2\nu x_1 + 2\nu x_2 - 2\nu = 0,$$

with solution $\boldsymbol{x}_\nu = \frac{\nu}{1+2\nu}(1,1)^{\mathrm{T}}$, for $\nu > 0$. Letting $\nu \to \infty$ we see that the sequence $\boldsymbol{x}_\nu$ converges to $(1/2, 1/2)^{\mathrm{T}}$ which is the unique optimal solution.

(1p) c) We note that the solution $\boldsymbol{x}_\nu$ fulfills $\nabla f(\boldsymbol{x}_\nu) + [2\nu h(\boldsymbol{x}_\nu)] \nabla h(\boldsymbol{x}_\nu) = \mathbf{0}$. So a Lagrange multiplier estimate comes from $\lambda_\nu := 2\nu h(\boldsymbol{x}_\nu)$. Insertion from b) yields $\lambda_\nu = \frac{-2\nu}{1+2\nu}$ which tends to $\lambda^* = -1$ when $\nu \to \infty$.

Question 5

(linear programming: existence of optimal solutions)

(2p) a) This is the first part of Theorem 8.10.

(1p) b) This is the second part of Theorem 8.10.

Question 6

(basic facts in optimization)

(1p) a) False. Any infeasible linear program will result in a phase I problem having a positive optimal value.

(1p) b) False. If f is strictly concave, and if at some $\boldsymbol{x} \in \mathbb{R}^n$, the Hessian matrix $\nabla^2 f(\boldsymbol{x})$ happens to be negative definite, then a search direction is well-defined, but it defines an ascent direction.

(1p) c) False. Consider the linear program to minimize x_2 subject to the constraints $0 \leq x_j \leq 4$, $j = 1, 2$, and the additional constraint that $x_1 + x_2 \leq 2$. This problem has the optimal solution set $X^* = \{ \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 \in [0, 2]; x_2 = 0 \}$. At the optimal solution $\boldsymbol{x}^* = (1, 0)^{\mathrm{T}}$, $x_1 + x_2 < 2$ holds. Believing that this means that the constraint $x_1 + x_2 \leq 2$ therefore is redundant results, however, in a grave mistake, as the new problem, having the constraints $0 \leq x_j \leq 4$, $j = 1, 2$, has the optimal set $X^*_{\text{new}} = \{ \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 \in [0, 4]; x_2 = 0 \}$.

Question 7

(Lagrangian duality)

(1p) a) At $\boldsymbol{x}^* = (3, 2)^{\mathrm{T}}$, the constraints $g_1(\boldsymbol{x}) := x_1 + x_2 \leq 5$ and $g_2(\boldsymbol{x}) := x_1 - 3 \leq 0$ are active. Introducing Lagrange multipliers for them, we study the equation

$$\nabla f(\boldsymbol{x}^*) + \mu_1 \nabla g_1(\boldsymbol{x}^*) + \mu_1 \nabla g_1(\boldsymbol{x}^*) = \mathbf{0}^2$$

at $\boldsymbol{x}^* = (3, 2)^{\mathrm{T}}$, that is,

$$\begin{pmatrix} -2 \\ -1 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence, $\mu_1 = \mu_2 = 1$.

As the KKT conditions are satisfied and the problem stated is convex, $\boldsymbol{x}^* = (3, 2)^{\mathrm{T}}$ is indeed a globally optimal solution.

(2p) b) With the Lagrangian $L(\boldsymbol{x}, \mu) := f(\boldsymbol{x}) + \mu(x_1 + x_2 - 5)$, we obtain the explicit Lagrangian dual function as follows:

First, the subproblem solution is that

$$x_1 = \begin{cases} 3, & \text{if } \mu \in [0, 2], \\ 5 - \mu, & \text{if } \mu \in [2, 5], \\ 0, & \text{if } \mu \in [2, \infty), \end{cases}$$

$$x_2 = \begin{cases} 3 - \mu, & \text{if } \mu \in [0, 3], \\ 0, & \text{if } \mu \in [3, \infty). \end{cases}$$

Hence,

$$q(\mu) \quad = \quad \begin{cases} -\frac{1}{2}\mu^2 + \mu + 2, & \text{if } \mu \in [0,2], \\ -\mu^2 + 3\mu, & \text{if } \mu \in [2,3], \\ -\frac{1}{2}\mu^2 + \frac{9}{2}, & \text{if } \mu \in [3,5], \\ -5\mu + 17, & \text{if } \mu \in [5,\infty), \end{cases}$$

which is to be maximized over non-negative values of μ.

Second, the derivative of q is

$$q'(\mu) \quad = \quad \begin{cases} -\mu + 1, & \text{if } \mu \in [0,2], \\ -2\mu + 3, & \text{if } \mu \in [2,3], \\ -\mu, & \text{if } \mu \in [3,5], \\ -5, & \text{if } \mu \in [5,\infty), \end{cases}$$

whence the optimal solution is found where q' changes sign, namely at $\mu^* = 1$.

This multiplier value is indeed the same as the one found as μ_1 in a). The corresponding primal solution then is $\boldsymbol{x}(\mu^*) = (3,2)^{\mathrm{T}}$. Further, $q(\mu^*) = 5/2$ equals the value of $f(\boldsymbol{x}^*)$, whence strong duality is fulfilled.