Chalmers/Gothenburg University
Mathematical Sciences

EXAM

TMA947/MMG621 OPTIMIZATION, BASIC COURSE

Date:	16–04–05
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:	Johannes Borgqvist (ankn. 5325)
Result announced:	16–04–15 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:

$$\begin{aligned} \text{minimize} \quad & z = 8x_1 + 3x_2 + 4x_3 + x_4, \\ \text{subject to} \quad & 2x_1 + x_2 + 3x_3 - x_4 = 5, \\ & x_1 + x_2 + 2x_3 - x_4 = 3, \\ & x_1, \quad x_2, \quad x_3, \quad x_4 \geq 0. \end{aligned}$$

Instead of trying to solve the problem using phase I and phase II simplex method separately, we could solve it in "one-shot". We consider the modified problem:

$$\begin{aligned} \text{minimize} \quad & z = 8x_1 + 3x_2 + 4x_3 + x_4 + My_1 + My_2, \\ \text{subject to} \quad & 2x_1 + x_2 + 3x_3 - x_4 + y_1 = 5, \\ & x_1 + x_2 + 2x_3 - x_4 + y_2 = 3, \\ & x_1, \quad x_2, \quad x_3, \quad x_4, \quad y_1, \quad y_2 \geq 0, \end{aligned}$$

where M is a very large but unspecified number such that $a + M > 0$ and $a - M < 0$ for all real number a.

(1p) a) Is the modified problem with M always feasible? Assume that the optimal objective value of the modified problem is bounded from below. If we solve the modified problem, what can we say about the feasibility and optimal objective value of the original problem, depending on the optimal values of y_1 and y_2 in the modified problem? Explain your answers.

(2p) b) Solve the modified problem with M using the simplex method, keeping M as a unspecified large number. If the problem can be solved to optimality, write down an optimal solution and objective value of the original problem.

Aid: Utilize the identity

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

Question 2

(true or false)

The below three claims should be assessed. Are they true or false, or is it impossible to say? Provide an answer, together with a short motivation.

(1p) a) Consider a standard LP problem, for which you apply the Simplex method. Suppose also that you have used Phase I of the simplex method and identified a basic feasible solution.

Claim: Then in Phase II you will be able to identify an optimal solution to the given problem.

(1p) b) Suppose that you are solving an unconstrained optimization problem in which you minimize a differentiable function f. Suppose further that at a given vector $\boldsymbol{x}$ you have generated a descent direction $\boldsymbol{p}$.

Claim: Then the Armijo rule will provide a positive, finite step length in which the objective function has a lower value of $f : \mathbb{R}^n \to \mathbb{R}$ than at $\boldsymbol{x}$.

(1p) c) Consider the problem of minimizing a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ over a bounded polyhedral set. Suppose further that we attack this problem by utilizing the Frank–Wolfe method. Suppose then that we have solved the linear subproblem of the algorithm.

Claim: Then the linearized objective function has an optimal value in the linear subproblem that is lower than or equal to the objective value at the current iteration.

(3p) Question 3

(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.

Establish Farkas' Lemma.

(3p) Question 4

(Frank–Wolfe)

Consider the problem to

$$\underset{x \in X}{\text{maximize}} \ f(\boldsymbol{x}). \tag{2}$$

Assume that X is a polyhedron and $f \in C^1$. Let $\bar{\boldsymbol{x}} \in X$ be a point to which the Frank–Wolfe algorithm converges within a finite number of iterations on the problem (2). Can we guarantee that $\bar{\boldsymbol{x}}$ is optimal? If not, which properties can we guarantee that $\bar{\boldsymbol{x}}$ has, and which additional requirements are necessary to guarantee that $\bar{\boldsymbol{x}}$ is an optimal solution to the problem (2)?

(3p) Question 5

(Lagrangian duality)

Consider the optimization problem

$$\begin{aligned} f^* := \underset{\boldsymbol{x}}{\text{infimum}} \quad & f(\boldsymbol{x}), \\ \text{subject to} \quad & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}^m, \\ & \boldsymbol{x} \in X. \end{aligned}$$

Let the Lagrange function be defined as $L(\boldsymbol{x}, \boldsymbol{\mu}) := f(\boldsymbol{x}) + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})$. Assume that $\boldsymbol{\mu}^*$ is a Lagrange multiplier. That is, $\boldsymbol{\mu}^* \geq \boldsymbol{0}^m$ and $\inf_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{\mu}^*) = f^*$. Show that $\boldsymbol{x}^*$ is optimal if and only if

$$\begin{aligned} \boldsymbol{x}^* \in X, \quad \boldsymbol{g}(\boldsymbol{x}^*) &\leq \boldsymbol{0}^m, \\ \boldsymbol{x}^* &\in \underset{\boldsymbol{x} \in X}{\arg\min} \ L(\boldsymbol{x}, \boldsymbol{\mu}^*), \\ \boldsymbol{\mu}_i^* g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, \ldots, m. \end{aligned}$$

(3p) Question 6

(integer programming modeling) Let a chessboard be a $n \times n$ grid with n being some integer. A queen can move any number of squares horizontally, vertically or diagonally. See Figure. 1 for an illustration of the possible moves of a queen.

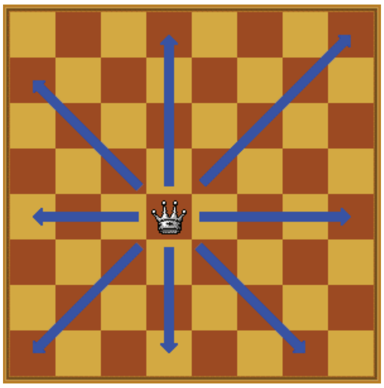

Figure 1: Possible moves of a queen. Source: `http://www.chess-poster.com`

For this problem, we can place an arbitrary number of queens on the chessboard. We are asked to find a configuration with the minimum number of queens so that

- each square either is occupied by a queen or can be attacked by a queen,
- no two queens can attack each other.

Formulate the problem to find the desirable configuration as an integer program.

(3p) Question 7

(gradient projection algorithm)

Consider the optimization problem to

$$\begin{aligned} \underset{x_1,x_2}{\text{minimize}} \quad & f(\boldsymbol{x}) := \tfrac{1}{2}(x_1 - 2)^2 + \tfrac{1}{2}(x_2 - \tfrac{3}{2})^2, \\ \text{subject to} \quad & \boldsymbol{x} \in X = \{(x_1, x_2)^{\mathrm{T}} \mid -1 \le x_i \le 1, \ \ i = 1, 2\}. \end{aligned}$$

We consider solving the problem using the gradient projection algorithm. Start with the initial point $\boldsymbol{x}^0 = (0, 0)^{\mathrm{T}}$. Perform one step of the gradient projection algorithm (so that you obtain the next iterate $\boldsymbol{x}^1$). Use the projection arc and perform exact minimization line search. That is, $\boldsymbol{x}^{k+1} = \mathrm{Proj}_X[\boldsymbol{x}^k + \alpha^k \boldsymbol{p}^k]$ for the appropriate search direction $\boldsymbol{p}^k$ and step size α^k for each iteration k. Is $\boldsymbol{x}^1$ optimal or not? Explain your answer.

Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date: 16–04–05
Examiner: Michael Patriksson

Question 1

(the simplex method)

(1p) a) The modified problem is always feasible by construction. For example, a feasible solution is $x_i = 0$ for $i = 1, 2, 3, 4$ and $y_1 = 5$ and $y_2 = 3$. Assuming that the modified problem has optimal objective value bounded from below, the modified problem always has finite optimal solution. Let $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ denote the x-part and y-part of the optimal solution, respectively. Depending on the value of $\boldsymbol{y}^*$, two cases are possible:

- At optimality, $y_1^* = y_2^* = 0$. In this case, the original problem is feasible. In addition, $\boldsymbol{x}^*$ is an optimal solution to the original problem. It is obvious that $\boldsymbol{x}^*$ is feasible to the original problem. If there were some $\tilde{x}$ feasible to the original problem with an objective value smaller than that of $\boldsymbol{x}^*$, then $\tilde{\boldsymbol{x}}$ together with $\boldsymbol{y}^* = \mathbf{0}$ form a better feasible solution to the modified problem. This contradicts the optimality of x^* and $\boldsymbol{y}^*$ for the modified problem.
- At optimality, at least one of y_1^* and y_2^* is positive. In this case, the original problem is infeasible. If a vector $\tilde{\boldsymbol{x}}$ were feasible to the original problem, then $\tilde{\boldsymbol{x}}$ together with $\boldsymbol{y} = \mathbf{0}$ result in a better feasible solution of the modified problem than $\boldsymbol{x}^*$ with $\boldsymbol{y}^*$ (cf. the property of M). This would contradicts the optimality of $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ for the modified problem.

(2p) b) We can start the simplex method with y_1 and y_2 being the basic variables. The non-basic variables are x_1, x_2, x_3 and x_4.

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{c}_N^{\mathrm{T}} = \begin{pmatrix} 8 & 3 & 4 & 1 \end{pmatrix}, \quad \boldsymbol{c}_B^{\mathrm{T}} = \begin{pmatrix} M & M \end{pmatrix}$$

$$N = \begin{pmatrix} 2 & 1 & 3 & -1 \\ 1 & 1 & 2 & -1 \end{pmatrix}, \quad x_B = B^{-1}b = \begin{pmatrix} 5 \\ 3 \end{pmatrix}.$$

The reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N = \begin{pmatrix} 8 - 3M & 3 - 2M & 4 - 5M & 1 + 2M \end{pmatrix}.$$

We choose the third non-basic variable (i.e., x_3) to enter the basis, because it has the most negative reduced cost. The corresponding search direction for the basic variables are $d_B = -B^{-1}N_3 = (-3, -2)^{\mathrm{T}}$. The minimum ratio test indicates that

$$2 = \operatorname{argmin}\{\frac{5}{3}, \frac{3}{2}\},$$

and hence the second basic variable (i.e., y_2) leaves the basis.

At iteration two, we have x_3 and y_1 being the basic variables. The non-basic variables are x_1, x_2, x_4 and y_2.

$$B = \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{3}{2} \end{pmatrix}, \quad \boldsymbol{c}_N^{\mathrm{T}} = \begin{pmatrix} 8 & 3 & 1 & M \end{pmatrix}, \quad \boldsymbol{c}_B^{\mathrm{T}} = \begin{pmatrix} 4 & M \end{pmatrix}$$

$$N = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 1 & 1 & -1 & 1 \end{pmatrix}.$$

The reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N = \begin{pmatrix} 6 - \frac{M}{2} & 1 + \frac{M}{2} & 3 - \frac{M}{2} & -2 + \frac{3M}{2} \end{pmatrix}.$$

We choose the third non-basic variable (i.e., x_4) to enter the basis. The corresponding search direction for the basic variables are $d_B = -B^{-1}N_3 = (\frac{1}{2}, -\frac{1}{2})^{\mathrm{T}}$. Therefore, the second basic variable (i.e., y_1) leaves the basis.

At iteration three, we have basic variables being x_3 and x_4. The non-basic variables are x_1, x_2, y_1 and y_2.

$$B = \begin{pmatrix} 3 & -1 \\ 2 & -1 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & -1 \\ 2 & -3 \end{pmatrix}, \quad \boldsymbol{c}_N^{\mathrm{T}} = \begin{pmatrix} 8 & 3 & M & M \end{pmatrix}, \quad \boldsymbol{c}_B^{\mathrm{T}} = \begin{pmatrix} 4 & 1 \end{pmatrix},$$

$$N = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \quad x_B = B^{-1}b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

The reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N = \begin{pmatrix} 3 & 4 & M - 6 & M + 7 \end{pmatrix}.$$

The reduced costs are all nonnegative. The simplex method terminates with optimal solution

$$x^* = (0, 0, 2, 1)^{\mathrm{T}}, \quad \boldsymbol{y}^* = (0, 0), \quad z^* = 9$$

As explained in part a), x^* is also an optimal solution to the original problem with objective value 9.

Question 2

(true or false)

(1p) a) Impossible to say, since the original problem may lack optimal solutions.

(1p) b) True—see Exercise 11.1.

(1p) c) Impossible to say, since the function f may not be convex.

(3p) Question 3

(optimality conditions)

This is Theorem 10.10.

(3p) Question 4

(Frank–Wolfe)

We can only guarantee that the point obtained is stationary. If f however is concave, then we establish that the point obtained is optimal.

(3p) Question 5

(Lagrangian duality)

This is Theorem 6.8.

(3p) Question 6

(integer programming modeling)

A suggested integer programming formulation is as follows: each square is labeled with an i nteger index (e.g., $1, \ldots, n^2$). For each square i, we define the neighborhood N_i to be the set of all indices of squares that can be attacked if a queen is placed at square i. For each i, we define a 0-1 binary decision variable $x_i \in \{0, 1\}$ such that a queen is placed at square i if and only if $x_i = 1$. Then, an integer program modeling t he desired queen configuration problem is

$$\begin{aligned}
\underset{\boldsymbol{x}}{\text{minimize}} \quad & \sum_{i=1}^{n^2} x_i \\
\text{subject to} \quad & x_i + \sum_{j \in N_i} x_j \geq 1, \quad i = 1, \ldots, n^2 \\
& (n^2 - 1)x_i + \sum_{j \in N_i} x_j \leq n^2 - 1, \quad i = 1, \ldots, n^2 \\
& x_i \in \{0, 1\}, \quad i = 1, \ldots, n^2.
\end{aligned}$$

In the model above, the first constraint specifies that for each square i either there is a queen or the square can be attacked by a queen in the neighborhood N_i. The second con straint specifies that if a queen is placed at square i, then no queen can be placed at a ny square in the neighborhood N_i (we can replace $n^2 - 1$ by any constant larger than tha t). The two constraints model exactly the conditions required by the queen configuration pr oblem.

(3p) Question 7

(gradient projection algorithm)

At $\boldsymbol{x}^0 = (0, 0)^{\mathrm{T}}$, the objective gradient vector is $\nabla f(\boldsymbol{x}^0) = (x_1 - 2, x_2 - \frac{3}{2})^{\mathrm{T}} = (-2, -\frac{3}{2})^{\mathrm{T}}$. Hence, the search direction is $\boldsymbol{p}^0 = -\nabla f(\boldsymbol{x}^0) = (2, \frac{3}{2})^{\mathrm{T}}$. Because of the form of the feasible set X (i.e., box constraints), projection on X can be expressed analytically. The projection arc is of the form (for $0 \leq \alpha^0 \leq 1$):

$$\mathrm{Proj}_X[\boldsymbol{x}^0 + \alpha^0 \boldsymbol{p}^0] = \begin{pmatrix} \min\{1, 0 + 2\alpha^0\} \\ \min\{1, 0 + \frac{3}{2}\alpha^0\} \end{pmatrix}.$$

Hence, the objective function (to be minimized) for exact line search is

$$\begin{aligned}
f^0(\alpha^0) \quad &:= \quad \tfrac{1}{2}(\min\{1, 2\alpha^0\} - 2)^2 + \tfrac{1}{2}(\min\{1, \tfrac{3}{2}\alpha^0\} - \tfrac{3}{2})^2 \\
&= \quad \begin{cases} \frac{1}{2}\Big(4(\alpha^0 - 1)^2 + \frac{9}{4}(\alpha^0 - 1)^2\Big) & 0 \leq \alpha^0 \leq \frac{1}{2} \\ \frac{1}{2}\Big(1 + \frac{9}{4}(\alpha^0 - 1)^2\Big) & \frac{1}{2} \leq \alpha^0 \leq \frac{2}{3} \\ \frac{5}{8} & \frac{2}{3} \leq \alpha^0 \leq 1 \end{cases}.
\end{aligned}$$

Minimizing f^0 with $0 \leq \alpha^0 \leq 1$ yields the minimizing α^0 to be greater than or equal to $2/3$. Hence, the next iterate is

$$\boldsymbol{x}^1 = \mathrm{Proj}_X[\boldsymbol{x}^0 + \alpha^0 \boldsymbol{p}^0] = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

It is claimed that $\boldsymbol{x}^1$ is an optimal solution. First, note that the objective gradient a t $\boldsymbol{x}^1 = (1,1)^{\mathrm{T}}$ is $\nabla f(\boldsymbol{x}^1) = (x_1 - 2, x_2 - \frac{3}{2})^{\mathrm{T}} = (-1, -\frac{1}{2})^{\mathrm{T}}$. At $\boldsymbol{x}^1$ the active constraints are $x_1 \leq 1$ and $x_2 \leq 1$ with constraint function gradients being $(1,0)^{\mathrm{T}}$ and $(0,1)^{\mathrm{T}}$, respectively. As a result, $-\nabla f(\boldsymbol{x}^1)$ is in the cone of the active constraint gradients. This implies that $\boldsymbol{x}^1$ is a KKT point. In addition, the optimization problem is convex with affine constraints. Hence, the KKT point $\boldsymbol{x}^1$ is indeed an optimal solution.