

**TMA947/MMG621
OPTIMIZATION, BASIC COURSE**

- Date:** 16-01-12
- Time:** Eklandagatan 86, morning, 8³⁰-13³⁰
- 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:** Edvin Wedin, tel. 0703-088304
- Result announced:** 16-02-02
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.*

(3p) Question 1

(the simplex method)

Consider the following linear program:

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

- (2p)** a) Solve the problem using phase I (so that you begin with a unit matrix as the first basis) and phase II of the simplex method. If the problem has an optimal solution, then present the optimal solution in both the original variables and in the variables used in the standard form. If the problem is unbounded, then use your calculations to find a direction of unboundedness in both the original variables and in the variables in the standard form.
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) Suppose that to the original problem we add a new variable x_4 and obtain the new problem to

$$\begin{aligned}
 \text{minimize} \quad & z = 2x_1 - x_2 + x_3 - \frac{1}{2}x_4, \\
 \text{subject to} \quad & x_1 + 3x_2 - x_3 + 8x_4 \leq 5, \\
 & -2x_1 + x_2 - 2x_3 - x_4 \leq -2, \\
 & x_1, \quad x_2, \quad x_3, \quad x_4 \geq 0.
 \end{aligned}$$

If the original problem has an optimal solution, explain how the optimal solution is affected by adding the new variable. If the original problem is unbounded, investigate if adding the new variable affects the unboundedness of the problem.

Note: Use your calculations from a) to answer the question.

(3p) Question 2

(Quadratic programming)

Consider the minimization of the quadratic function $f(\mathbf{y}) := \frac{1}{2}\mathbf{y}^T \mathbf{A}\mathbf{y} + \mathbf{b}^T \mathbf{y} + c$ subject to the constraints $\mathbf{y} \geq \mathbf{0}^n$, where \mathbf{A} is symmetric and positive semi-definite. Show that the three conditions $\mathbf{A}\mathbf{x} + \mathbf{b} \geq \mathbf{0}^n$, $\mathbf{x} \geq \mathbf{0}^n$, and $\mathbf{x}^T \mathbf{A}\mathbf{x} + \mathbf{b}^T \mathbf{x} = \mathbf{0}$ are necessary and sufficient to characterize \mathbf{x} as a minimum.

(3p) Question 3(characterization of convexity in C^1)

Let $f \in C^1$ on an open convex set S . Establish the following characterization of the convexity of f on S :

f is convex on $S \iff f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x})$, for all $\mathbf{x}, \mathbf{y} \in S$.

Question 4

(true or false claims in optimization)

For each of the following three claims, your task is to decide whether it is true or false. Motivate your answers clearly.

(1p) a) If, in the solution of a minimization problem of the linear programming type, a current non-degenerate BFS (basic feasible solution) has a non-negative vector of reduced costs, then that BFS corresponds to an optimal extreme point solution in the problem.

(1p) b) If, in the solution of an unconstrained optimization problem of the form

$$\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} f(\mathbf{x}),$$

with $f : \mathbb{R}^n \rightarrow \mathbb{R}$ being in C^1 , you have found a vector $\bar{\mathbf{x}}$ with $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}^n$, then $\bar{\mathbf{x}}$ is an optimal solution to the problem.

(1p) c) Suppose you have solved the problem to

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

where $X = \{ \mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \}$, the functions f and g_i , $i = 1, \dots, m$, are continuous, and the vector $\bar{\mathbf{x}}$ is an optimal solution. Suppose further that for some $j = 1, \dots, m$, $g_j(\bar{\mathbf{x}}) < 0$. Then, that constraint is *redundant*, that is, if that constraint is removed, then the remaining problem with $m - 1$ constraints also has $\bar{\mathbf{x}}$ as an optimal solution.

(3p) Question 5

(KKT conditions)

Consider the problem

$$\begin{aligned} & \text{minimize} && f(\mathbf{x}) := x_1 \\ & \text{subject to} && x_1^2 + x_2^2 \leq 2 \\ & && (x_1 - 2)^2 + (x_2 - 2)^2 \leq 2 \end{aligned}$$

- (1p)** Establish theoretically or graphically that $\mathbf{x}^* = (1, 1)^T$ is the unique globally optimal solution.
- (2p)** Determine if the KKT conditions are satisfied at \mathbf{x}^* . If they are not, explain why, and relate your explanation to the known results on necessary and sufficient optimality conditions.

(3p) Question 6

(Frank-Wolfe algorithm)

Consider the problem

$$\begin{aligned} & \text{minimize}_{x_1, x_2} && f(\mathbf{x}) := (x_1 \quad x_2) \begin{pmatrix} 6 & 2 \\ 2 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - (52 \quad 34) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ & \text{subject to} && x_1 + 2x_2 \leq 4 \\ & && x_1 + x_2 \leq 3 \\ & && 2x_1 \leq 5 \\ & && x_1 \geq 0 \\ & && x_2 \geq 0 \end{aligned} \tag{1}$$

Solve problem (1) with the Frank-Wolfe algorithm. Start with initial guess $\mathbf{x}^{(0)} = (x_1, x_2)^T = (0, 0)^T$. Use exact minimization for line search. If necessary, you are allowed to carry out the calculations approximately with two digits of accuracy.

Hint: You may find it helpful to analyze the problem and the algorithm progress in picture, but this should be augmented with rigorous analysis.

(3p) **Question 7**

(LP duality)

Consider the problem

$$\begin{aligned} & \underset{\mathbf{x}}{\text{maximize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \inf_{\mathbf{y} \in P} \mathbf{y}^T \mathbf{x} \geq d \\ & && \mathbf{x} \geq \mathbf{0}^n, \end{aligned}$$

where the problem data are $\mathbf{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$, $P = \{\mathbf{y} \mid \mathbf{A}\mathbf{y} \geq \mathbf{b}, \mathbf{y} \geq \mathbf{0}^n\}$ with $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$. It is assumed that P is nonempty and bounded. Show that the problem can be written as a linear program as

$$\begin{aligned} & \underset{\mathbf{x}, \mathbf{z}}{\text{maximize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \mathbf{A}^T \mathbf{z} \leq \mathbf{x} \\ & && \mathbf{b}^T \mathbf{z} \geq d \\ & && \mathbf{x} \geq \mathbf{0}^n, \mathbf{z} \geq \mathbf{0}^m. \end{aligned}$$

Chalmers/GU
Mathematics

EXAM SOLUTION

**TMA947/MMG621
NONLINEAR OPTIMISATION**

Date: 16-01-12

Examiner: Michael Patriksson

Question 1

(the simplex method)

- (2p) a) We first rewrite the problem in standard form. We introduce slack variables s_1 and s_2 . Consider the following linear program:

$$\begin{aligned} \text{minimize} \quad & z = 2x_1 - x_2 + x_3 \\ \text{subject to} \quad & x_1 + 3x_2 - x_3 + s_1 = 5, \\ & 2x_1 - x_2 + 2x_3 - s_2 = 2, \\ & x_1, \quad x_2, \quad x_3, \quad s_1, \quad s_2 \geq 0. \end{aligned}$$

Phase I

We introduce an artificial variable a and formulate our Phase I problem.

$$\begin{aligned} \text{minimize} \quad & z = a \\ \text{subject to} \quad & x_1 + 3x_2 - x_3 + s_1 = 5, \\ & 2x_1 - x_2 + 2x_3 - s_2 + a = 2, \\ & x_1, \quad x_2, \quad x_3, \quad s_1, \quad s_2, \quad a \geq 0. \end{aligned}$$

We now have a starting basis (s_1, a) . Calculating the reduced costs we obtain $\tilde{\mathbf{c}}_N = (-2, 1, -2, 1)^T$, meaning that x_1 or x_3 should enter the basis. We choose x_3 . From the minimum ratio test, we get that a should leave the basis. This concludes Phase I and we now have the basis (s_1, x_3) .

Phase II

Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (1, -\frac{1}{2}, \frac{1}{2})^T$. meaning that x_2 should enter the basis. From the minimum ratio test, we get that the outgoing variable is s_1 . Updating the basis we now have (x_2, x_3) in the basis.

Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (\frac{7}{5}, \frac{1}{5}, \frac{2}{5})^T \geq 0$, meaning that the current basis is optimal. The optimal solution is thus

$$(x_1, x_2, x_3, s_1, s_2)^T = (0, \frac{12}{5}, \frac{11}{5}, 0, 0, 0)^T,$$

which in the original variables means $(x_1, x_2, x_3)^T = (0, \frac{12}{5}, \frac{11}{5})^T$ with optimal objective value $f^* = -\frac{1}{5}$.

- (1p) b) Calculating the reduced costs of the modified problem for the optimal basis of the original problem, we obtain $\tilde{\mathbf{c}}_N = (\frac{7}{5}, \frac{1}{5}, \frac{2}{5}, \frac{7}{10})^T \geq 0$ meaning that the the optimal basis from the original problem gives the optimal solution of the modified problem $(x_1, x_2, x_3, x_4)^T = (0, \frac{12}{5}, \frac{11}{5}, 0)^T$ with optimal objective value $f^* = -\frac{1}{5}$.

Question 2

(Quadratic programming)

Since the objective function is convex (i.e., Hessian matrix \mathbf{A} is symmetric positive semidefinite) and the constraints are affine, the KKT conditions are both necessary and sufficient for optimality. Therefore, a point \mathbf{x} is a minimum if and only if there exists a vector $\boldsymbol{\mu} \in \mathbb{R}^n$ (Lagrangian multipliers) such that

$$\begin{aligned} \mathbf{Ax} + \mathbf{b} &= \boldsymbol{\mu} \\ \boldsymbol{\mu} &\geq \mathbf{0}^n \\ \mathbf{x} &\geq \mathbf{0}^n \\ \boldsymbol{\mu}_i \mathbf{x}_i &= 0, \quad \forall i = 1, \dots, n. \end{aligned}$$

Eliminating $\boldsymbol{\mu}$, the above conditions are equivalent to

$$\begin{aligned} \mathbf{Ax} + \mathbf{b} &\geq \mathbf{0}^n \\ \mathbf{x} &\geq \mathbf{0}^n \\ (\mathbf{Ax} + \mathbf{b})_i \mathbf{x}_i &= 0, \quad \forall i = 1, \dots, n. \end{aligned}$$

These are in turn equivalent to

$$\begin{aligned} \mathbf{Ax} + \mathbf{b} &\geq \mathbf{0}^n \\ \mathbf{x} &\geq \mathbf{0}^n \\ \mathbf{x}^\top \mathbf{Ax} + \mathbf{b}^\top \mathbf{x} &= 0. \end{aligned}$$

Question 3(characterization of convexity in C^1)

This is Theorem 3.61 (a) in the textbook.

Question 4

(true or false claims in optimization)

- (1p) a) The claim is true, as stated by Proposition 9.1 in the textbook.

- (1p) b) The claim is false. The point $\bar{\mathbf{x}}$ with $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}^n$ can also be a local maximum or saddle point.
- (1p) c) The claim is false. Consider the problem with one decision variable. $f(x) = \min\{0, -x\}$ and $g(x) = x$. The point $\bar{x} = -1$ is a constrained minimum and $g(\bar{x}) = -1 < 0$. However, removing the constraint $g(x) \leq 0$ will result in a problem whose objective value is unbounded from below.

Question 5

(KKT conditions)

- (1p) a) The point $\mathbf{x}^* = (1, 1)^T$ is the only feasible point and hence it must be the unique global minimum.
- (2p) b) Let $g_1(\mathbf{x}) := x_1^2 + x_2^2 - 2$ and $g_2(\mathbf{x}) := (x_1 - 2)^2 + (x_2 - 2)^2 - 2$. At $\mathbf{x}^* = (1, 1)^T$, both $g_1(\mathbf{x}^*) = 0$ and $g_2(\mathbf{x}^*) = 0$. That is, both inequality constraints are active. Also, it holds that

$$\nabla f(\mathbf{x}^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla g_1(\mathbf{x}^*) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \nabla g_2(\mathbf{x}^*) = -\begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Therefore, the equality (as part of KKT conditions)

$$-\nabla f(\mathbf{x}^*) = \mu_1 \nabla g_1(\mathbf{x}^*) + \mu_2 \nabla g_2(\mathbf{x}^*), \quad \mu_1 \geq 0, \mu_2 \geq 0$$

cannot hold. Hence, the KKT conditions are not satisfied. As a result, the KKT conditions are not necessary for optimality since \mathbf{x}^* is a minimum but not a KKT point. This does not contradict any result regarding the necessity of the KKT conditions. For instance, $\nabla g_1(\mathbf{x}^*)$ and $\nabla g_2(\mathbf{x}^*)$ are not linearly independent, and hence the LICQ constraint qualification does not hold. On the other hand, since the problem is convex, KKT points (if exist) are global optimal solutions.

Question 6

(Frank-Wolfe algorithm)

Figure 1 shows the feasible set of the problem (i.e., the polyhedron with thick black boundary lines) and some contours of the objective function. The optimal solution is denoted by x^* (i.e., the red dot in the figure).

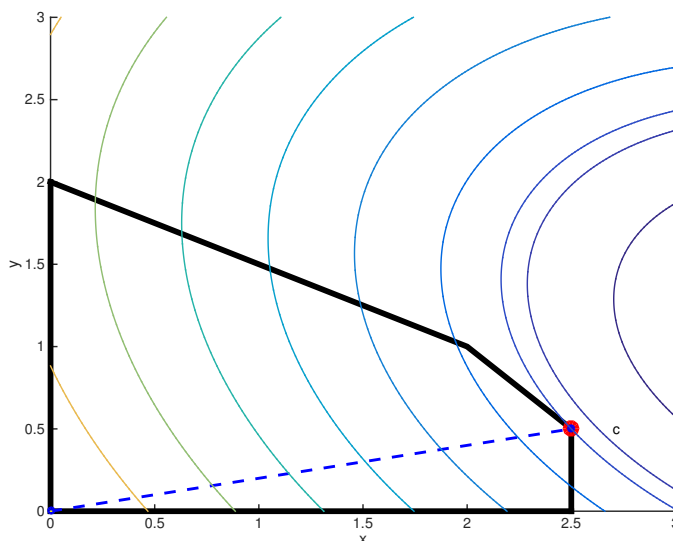


Figure 1: Illustration of the Frank-Wolfe algorithm. The feasible set is a polyhedron with boundary denoted by the thick black lines. Some contours of the objective function are shown. The optimal solution $x^* = (2.5, 0.5)^T$.

The details of the algorithm steps are as follows. Let X denote the feasible set. Let $f(x_1, x_2)$ denote the objective function. For any given iterate $x^{(k)} = (x_1^{(k)}, x_2^{(k)})^T$. The objective function gradient vector is

$$\nabla f(x_1^{(k)}, x_2^{(k)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix}.$$

The search direction problem is

$$\underset{x \in X}{\text{minimize}} \quad \nabla f(x_1^{(k)}, x_2^{(k)})^T x. \quad (1)$$

If $\min_{x \in X} \nabla f(x_1^{(k)}, x_2^{(k)})^T x \geq \nabla f(x_1^{(k)}, x_2^{(k)})^T x^{(k)}$, then by optimality conditions (for minimizing a convex function over a convex feasible set) $x^{(k)}$ is optimal. Otherwise, let $y^{(k)}$ denote an optimal solution to the search direction problem. Then the exact minimization line search problem can be expressed into

$$\underset{\alpha \in [0,1]}{\text{minimize}} \quad f(\alpha x^{(k)} + (1 - \alpha)y^{(k)}) \iff \underset{\alpha \in [0,1]}{\text{minimize}} \quad g\alpha^2 + h\alpha,$$

where

$$\begin{aligned} g &= (x^{(k)} - y^{(k)})^T \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} (x^{(k)} - y^{(k)}) \\ h &= (x^{(k)} - y^{(k)})^T \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} y^{(k)} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right). \end{aligned} \quad (2)$$

The minimizing value of α , denoted by $\alpha^{(k)}$, can be found using the optimality condition to be

$$\alpha^{(k)} = \begin{cases} 0 & \text{if } -\frac{h}{2g} < 0 \\ -\frac{h}{2g} & \text{if } 0 \leq -\frac{h}{2g} \leq 1. \\ 1 & \text{if } -\frac{h}{2g} > 1 \end{cases} \quad (3)$$

The iterate update formula is

$$x^{(k+1)} = \alpha^{(k)} x^{(k)} + (1 - \alpha^{(k)}) y^{(k)}. \quad (4)$$

Now we begin applying the Frank-Wolfe algorithm. At the first iteration with $x^{(0)} = (0, 0)$, the objective function gradient is

$$\nabla f(x_1^{(0)}, x_2^{(0)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \begin{bmatrix} -52 \\ -34 \end{bmatrix}.$$

To solve the search direction problem in (1), it is sufficient to restrict the feasible set to the set of all extreme points. That is,

$$\underset{x \in V}{\text{minimize}} \quad \nabla f(x_1^{(0)}, x_2^{(0)})^T x, \quad (5)$$

where V is the set of all extreme points defined as

$$V = \left\{ (0, 0)^T, (0, 2)^T, (2, 1)^T, (2.5, 0.5)^T, (2.5, 0)^T \right\}.$$

This amounts to finding the minimum among five numbers: 0, -68 , -138 , -147 , -130 . The result is that $y^{(0)} = (2.5, 0.5)^T$. Applying the formula in (2) yields

$$\begin{aligned} g &= \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 0.5 \end{bmatrix} \right)^T \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 0.5 \end{bmatrix} \right) = 44.75 \\ h &= \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 0.5 \end{bmatrix} \right)^T \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} 2.5 \\ 0.5 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right) = 57.5 \end{aligned}$$

According to (3), $\alpha^{(0)} = 0$. Hence, by (4)

$$x^{(1)} = y^{(0)} = (2.5, 0.5)^T.$$

This is shown in Figure 1.

At the next iteration with $x^{(1)} = (2.5, 0.5)^T$, we have

$$\nabla f(x_1^{(1)}, x_2^{(1)}) = \begin{bmatrix} -20 \\ -15 \end{bmatrix}.$$

Solving (5) leads to $y^{(1)} = x^{(1)} = (2.5, 0.5)^T$. Thus, it holds that

$$\min_{x \in X} \nabla f(x_1^{(1)}, x_2^{(1)})^T x \geq \nabla f(x_1^{(1)}, x_2^{(1)})^T x^{(1)}.$$

By optimality conditions, $x^{(1)} = (2.5, 0.5)^T$ is the optimal solution to our problem.

Question 7

(LP duality)

Since $P = \{\mathbf{y} \mid \mathbf{A}\mathbf{y} \geq \mathbf{b}, \mathbf{y} \geq \mathbf{0}^n\}$ is assumed to be nonempty and bounded, strong duality implies that, for any fixed \mathbf{x} , the minimum objective value of

$$\begin{aligned} \inf_{\mathbf{y}} \quad & \mathbf{y}^T \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{y} \geq \mathbf{b} \\ & \mathbf{y} \geq \mathbf{0}^n \end{aligned}$$

is the same as the maximum objective value of

$$\begin{aligned} \sup_{\mathbf{z}} \quad & \mathbf{b}^T \mathbf{z} \\ \text{subject to} \quad & \mathbf{A}^T \mathbf{z} \leq \mathbf{x} \\ & \mathbf{z} \geq \mathbf{0}^m. \end{aligned} \tag{1}$$

Substituting (1) into the original problem in the statement of Problem 7 results in

$$\begin{aligned} \text{maximize}_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{subject to} \quad & \sup_{\mathbf{z}} \mathbf{b}^T \mathbf{z} \geq d \\ & \mathbf{A}^T \mathbf{z} \leq \mathbf{x} \\ & \mathbf{x} \geq \mathbf{0}^n, \mathbf{z} \geq \mathbf{0}^m. \end{aligned}$$

This problem is equivalent to the second problem in the statement of Problem 7.