Chalmers/Gothenburg University
Mathematical Sciences

EXAM

TMA947/MMG621
OPTIMIZATION, BASIC COURSE

Date:	15–04–14
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:	Anders Martinsson (0703-088304)
Result announced:	15–05–06 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:

$$\begin{aligned}
\text{maximize} \quad & z = 2x_1 + x_2, \\
\text{subject to} \quad & -x_1 + x_2 \leq 1, \\
& -x_1 + 2x_2 \geq -2. \\
& x_1, \quad x_2 \geq 0.
\end{aligned}$$

(2p) a) Solve the 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) If an optimal solution exists, then use your calculations to decide whether it is unique or not. If the problem is unbounded, then use your calculations to specify a direction of unboundedness of the objective value.

(3p) Question 2

(consistency of linear systems)

Consider the following system of linear inequalities:

$$\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}.$$

Suppose that this system has at least one solution. Let d be a given scalar. Use linear programming duality to establish the equivalence of the following two statements:

(a) Every solution $\boldsymbol{x}$ to the system $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ satisfies $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq d$.

(b) There exists some vector $\boldsymbol{y} \geq \boldsymbol{0}^n$ such that $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{c}$ and $\boldsymbol{b}^T\boldsymbol{y} \leq d$.

(3p) Question 3

(global optimality conditions)

The following result characterizes every optimal primal and dual solution. It is however applicable only in the presence of Lagrange multipliers; in other words, the below system (1) is consistent if and only if there exists a Lagrange multiplier vector and there is no duality gap.

THEOREM 1 (global optimality conditions in the absence of a duality gap) *The vector* $(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ *is a pair of primal optimal solution and Lagrange multiplier vector if and only if*

$$\boldsymbol{\mu}^* \geq \mathbf{0}^m, \qquad \text{(Dual feasibility)} \tag{1a}$$

$$\boldsymbol{x}^* \in \arg\min_{x \in X} L(\boldsymbol{x}, \boldsymbol{\mu}^*), \qquad \text{(Lagrangian optimality)} \tag{1b}$$

$$\boldsymbol{x}^* \in X, \quad \boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}^m, \qquad \text{(Primal feasibility)} \tag{1c}$$

$$\mu_i^* g_i(\boldsymbol{x}^*) = 0, \quad i = 1, \dots, m. \qquad \text{(Complementary slackness)} \tag{1d}$$

Establish this theorem.

(3p) Question 4

(modelling)

A company can produce two products, A and B. To produce one unit of product A takes two hours, while the corresponding time for product B is three hours. The profit for each unit of product A is 200 kr and for product B 400 kr. There are 40 hours of production time available. However, by paying a fixed cost of 1200 kr another 8 extra hours of production can be used. If more than 10 units of product A is produced, then at least five units of product B must be produced. It is only possible to produce integer number of products.

Formulate a linear integer program (a linear objective function, linear constraints, and integrality restrictions on the variables) to determine the optimal production plan to maximize the profit.

Question 5

(true or false)

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

(1p) a) Consider a convex function $f : \mathbb{R}^n \to \mathbb{R}$. Suppose that at some vector $\boldsymbol{x}$ the directional derivative of f in the direction of a given vector $\boldsymbol{p} \in \mathbb{R}^n$ is non-negative.

Claim: The vector $\boldsymbol{x}$ is a minimizer of f over $\mathbb{R}^n$.

(1p) b) Consider solving a linear program (call it "P") through the process of utilizing "phase I" and "phase II" of the Simplex method. Suppose that the optimal value in the phase I-problem is zero.

Claim: There exists an optimal solution to the linear program P.

(1p) c) Consider the problem of minimizing a differentiable convex 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 having solved the linear subproblem of the algorithm we find that the linearized objective function has an optimal value in the linear subproblem that is equal to the objective value at the current iteration.

Claim: Then the last iterate of the Frank–Wolfe method is optimal in the problem.

(3p)
Question 6

(nonlinear programming) Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}), \\ \text{subject to} \quad & \boldsymbol{x} \in X, \end{aligned}$$

where f is in C^1 and where

$$X = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \sum_{j=1}^{n} x_j = r; \quad x_j \geq 0, \; j = 1, \ldots, n \right\},$$

where $r > 0$. Suppose that $\boldsymbol{x}^*$ is a local optimum in this problem.

Show that

$$x_j^* > 0 \quad \Longrightarrow \quad \frac{\partial f(\boldsymbol{x}^*)}{\partial x_j} \leq \frac{\partial f(\boldsymbol{x}^*)}{\partial x_i}, \qquad i, j = 1, \dots, n,$$

that is, that variables with non-zero optimal values have the same (and minimal) partial derivatives.

(3p) Question 7

(gradient projection algorithm) Solve the following problem, utilizing the gradient projection method:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(\boldsymbol{x}) := 3x_1^2 - 2x_1x_2 + 2x_2^2, \\ \text{subject to} \quad & 0 \leq x_1 \leq 2, \\ & -3 \leq x_2 \leq -1. \end{aligned}$$

Initiate the algorithm at $\boldsymbol{x}^0 = (1, -2)^{\mathrm{T}}$, and utilize the Armijo criterion to determine the step length. Apply the gradient projection method for at most three iterations. When the algorithm terminates, either because of the iteration limit or a termination criterion being met, can you show whether the final iterate is indeed an optimal solution for this problem?

The Armijo criterion for step length determination is as follows: let f denote the objective function, and X denote the feasible set. Accept as step length (for iteration k) $\alpha^k = \bar{\alpha}\beta^i$, where i is the first nonnegative integer (starting with $0, 1, \dots$) such that

$$f(\mathrm{Proj}_X[\boldsymbol{x}^k - \bar{\alpha}\beta^i \nabla f(\boldsymbol{x}^k)]) \leq f(\boldsymbol{x}^k) + \mu \nabla f(\boldsymbol{x}^k)^{\mathrm{T}} \left(\mathrm{Proj}_X[\boldsymbol{x}^k - \bar{\alpha}\beta^i \nabla f(\boldsymbol{x}^k)] - \boldsymbol{x}^k \right), \tag{1}$$

where $\bar{\alpha} = 1$, $\beta = 0.5$ and $\mu = 0.2$. To clarify, in (1) $\boldsymbol{x}^k$ should be interpreted as the iterate at iteration k, while β^i should be interpreted as β to the i-th power.

Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 15–04–14

Examiner: Michael Patriksson

Question 1

(the simplex method)

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

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

In phase I the starting basis is $(s_1, s_2)^{\mathrm{T}}$. Calculating the reduced costs for the non-basic variables x_1, x_2 we obtain $\tilde{\mathbf{c}}_N = (-2, -1)^{\mathrm{T}}$, meaning that x_1 enters the basis. From the minimum ratio test, we get that s_2 leaves the basis.

Updating the basis we now have $(s_1, x_1)^{\mathrm{T}}$ in the basis. Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (2, -5)^{\mathrm{T}}$. meaning that x_2 enters the basis. From the minimum ratio test we get $\mathbf{B}^{-1}\mathbf{N}_2 = (-1, -2)^{\mathrm{T}} < \mathbf{0}$, meaning that the problem is unbounded.

(1p) b) A direction of unboudness is $\mathbf{l}(\mu) = (2, 0, 3, 0)^{\mathrm{T}} + \mu(2, 1, 1, 0)^{\mathrm{T}}$, $\mu \geq 0$.

(3p) Question 2

(consistency of linear systems)

Consider to linear program to

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & -\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \end{aligned} \tag{1}$$

and its standard form equivalent

$$\begin{aligned} \underset{x^+, x^-, s}{\text{minimize}} \quad & -\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^+ + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^-, \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x}^+ - \boldsymbol{A}\boldsymbol{x}^- + \boldsymbol{s} = \boldsymbol{b}, \\ & \boldsymbol{x}^+ \geq \mathbf{0}, \ \boldsymbol{x}^- \geq \mathbf{0}, \ \boldsymbol{s} \geq \mathbf{0}. \end{aligned} \tag{2}$$

The dual of (2) is to

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{p}, \\ \text{subject to} \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{p} \leq -\boldsymbol{c}, \\ & -\boldsymbol{A}^{\mathrm{T}}\boldsymbol{p} \leq \boldsymbol{c}, \\ & \boldsymbol{p} \leq \boldsymbol{0}, \end{aligned} \tag{3}$$

and (3) is equivalent to

$$\begin{aligned} \underset{y}{\text{maximize}} \quad & -\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}, \\ \text{subject to} \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{c}, \\ & \boldsymbol{y} \geq \boldsymbol{0}. \end{aligned} \tag{4}$$

If statement (a) holds, then the optimal objective value of (1) and (2) are bounded from below by $-d$. Hence, by there exists an optimal solution to the dual of (2), which is (3). Consequently, by strong duality (cf. Theorem 10.6) the optimal objective values of (3) and (4) are equal to that of (2), which is bounded from below by $-d$. This implies that, for (4), there exists a vector $\boldsymbol{y} \geq \boldsymbol{0}$ such that $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{c}$ and $-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \geq -\boldsymbol{d}$ (i.e., $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{d}$). This statement is the same as (b).

Conversely, if (b) holds then (3) has at least one feasible solution with objective value bounded from below by $-d$. Hence, by weak duality (cf. Theorem 10.4) every $\boldsymbol{x}$ feasible in (1) (i.e., $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$) must satisfy $-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \geq -d$. This implies statement (a).

(3p) Question 3

(global optimality conditions)

This is Theorem 6.8.

(3p) Question 4

(modelling)

The decision variables are:

x_A = number of units of product A produced

x_B = number of units of product B produced

y_1 = 1, if additional time is used, 0 otherwise

y_2 = 1, if more than ten units of product A is produced, 0 otherwise

Based on these definitions, the model is as follows:

$$\begin{aligned} \text{maximize} \quad & 200x_A + 400x_B - 1200y_1 \\ \text{subject to} \quad & 2x_A + 3x_B \le 40 + 8y_1, \\ & 100y_2 \ge x_A - 10, \\ & x_B \ge 5y_2, \\ & x_A, x_B \ge 0, \text{integer} \\ & y_1, y_2 \in \{0, 1\} \end{aligned}$$

The program is linear with integer variables.

Question 5

(true or false)

(1p) a) False. The directional derivative must be non-negative in *all* directions $\boldsymbol{p}$.

(1p) b) False. The problem is feasible but may have an unbounded solution.

(1p) c) True. This is a consequence of Theorem 4.23.

(3p) Question 6

(nonlinear programming)

Letting μ denote the Lagrange multiplier for the equality constraint, and $\boldsymbol{\lambda} \in \mathbb{R}^n_+$ denote the vector of multipliers for the sign constraints, we obtain the Lagrangian

$$L(\boldsymbol{x}, \mu, \boldsymbol{\lambda}) := f(\boldsymbol{x}) + \mu \left(\sum_{j=1}^{n} x_j - r \right) - \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{x}.$$

Consider the optimality condition for x_j:

$$\frac{\partial L(\boldsymbol{x}, \mu, \boldsymbol{\lambda})}{\partial x_j} = \frac{\partial f(\boldsymbol{x})}{\partial x_j} - \mu - \lambda_j = 0, \quad j = 1, \dots, n.$$

Further, we have that $\lambda_j^* x_j^* = 0$, by complementarity. If $x_j^* > 0$ then $\lambda_j^* = 0$, and hence $\frac{\partial f(x^*)}{\partial x_j} = \mu^*$ (hence a common partial derivative for *all* positive variables), while if $x_j^* = 0$ then $\frac{\partial f(x^*)}{\partial x_j} = \mu^* + \lambda_j^*$, which may be larger.

(3p) Question 7

(gradient projection algorithm)

Denote the objective function by $f(x_1, x_2) := 3x_1^2 - 2x_1x_2 + 2x_2^2$, and the (box) feasible set by X. Then, $\nabla f(\boldsymbol{x}) = (6x_1 - 2x_2, -2x_1 + 4x_2)^{\mathrm{T}}$. At the initial point $x^0 = (1, -2)^{\mathrm{T}}$, the gradient is $\nabla f(x^0) = (10, -10)^{\mathrm{T}}$. To determine step length α^0, we apply the Armijo criterion supplied. We first try $\alpha^0 = \bar{\alpha} = 1$ (as $\beta^0 = 1$). Note that

$$\mathrm{Proj}_X[\boldsymbol{x}^0 - \nabla f(\boldsymbol{x}^0)] = \mathrm{Proj}_X\left[\begin{pmatrix} 1-10 \\ -2+10 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\mathrm{Proj}_X[\boldsymbol{x}^0 - \nabla f(\boldsymbol{x}^0)] - \boldsymbol{x}^0 = \begin{pmatrix} 0 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$f\left(\mathrm{Proj}_X[\boldsymbol{x}^k - \bar{\alpha}\beta^i \nabla f(\boldsymbol{x}^k)]\right) = 3 \cdot 0 + 2 \cdot 0 - 2 \cdot (-1)^2 = 2$$

$$f(\boldsymbol{x}^0) = 3 \cdot 1^2 - 2 \cdot 1 \cdot (-2) + 2 \cdot (-2)^2 = 15$$

$$\nabla f(\boldsymbol{x}^0)^{\mathrm{T}} \left(\mathrm{Proj}_X[\boldsymbol{x}^0 - \nabla f(\boldsymbol{x}^0)] - \boldsymbol{x}^0\right) = \begin{pmatrix} 10 & -10 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -20.$$

Hence, the Armijo criterion is satisfied, as $2 \leq 15 + 0.2 \cdot (-20) = 11$. Thus, $\alpha^0 = \bar{\alpha} = 1$, and iterate $\boldsymbol{x}^1 = \mathrm{Proj}_X[\boldsymbol{x}^0 - \nabla f(\boldsymbol{x}^0)] = (0, -1)^{\mathrm{T}}$.

For the next iteration, we have $\nabla f(\boldsymbol{x}^1) = (2, -4)^{\mathrm{T}}$. Hence,

$$\boldsymbol{x}^1 - \alpha \nabla f(\boldsymbol{x}^1) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \begin{pmatrix} -2\alpha \\ 4\alpha \end{pmatrix} = \begin{pmatrix} -2\alpha \\ -1 + 4\alpha \end{pmatrix}.$$

As a result,

$$\mathrm{Proj}_X[\boldsymbol{x}^1 - \alpha \nabla f(\boldsymbol{x}^1)] = \begin{pmatrix} \max\{0, -2\alpha\} \\ \min\{-1, -1 + 4\alpha\} \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \boldsymbol{x}^1, \quad \forall \alpha > 0,$$

and hence $\boldsymbol{x}^1$ is a stationary point (KKT point). The gradient projection algorithm terminates because the termination criterion is met. Notice that the fact that $\boldsymbol{x}^1$ is a stationary point can also be understood graphically, as $-\nabla f(\boldsymbol{x}^1)$ lies in the cone generated by the normal vectors of the two active constraints ($x_1 \geq 0$ and $x_2 \leq -1$).

Finally, since f is convex (which can be verified by computing the Hessian) and X is convex, the stationary point $\boldsymbol{x}^1$ is also optimal. This can be established via Theorem 4.23 together with (4.18), or Theorem 5.49 in the text.
