Chalmers/GU
Mathematics

EXAM

TMA947/MMG621
OPTIMIZATION, BASIC COURSE

Date:	15–01–13
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:	Magnus Önnheim (0703-088304)
Result announced:	15–01–29
	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:

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

(2p) a) Solve the problem for $\alpha = -1$ 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) Find the values of α such that the optimal solution from a) is optimal.

(3p) Question 2

(convexity)

Consider the problem to minimize a convex function $f : \mathbb{R}^n \to \mathbb{R}$ over a nonempty, closed and convex set S. Suppose further that $\boldsymbol{x}^*$ is a locally optimal solution to this problem. Is it then globally optimal in the problem? Argue in detail.

Question 3

(KKT optimality conditions)

Consider the problem to project (according to the standard Euclidean distance) the vector $\boldsymbol{z} = (2, 3/2)^{\mathrm{T}}$ onto the set S specified by the constraints that $x_j \geq 0$ for $j = 1, 2$, and that $x_1 + x_2 \leq 3/2$.

(1p) a) Describe the appropriate optimization problem to be solved in order to find this projection, and establish that it is a convex optimization problem. (Note: Use the square of the Euclidian distance as objective function.)

(1p) b) State the KKT conditions corresponding to a feasible vector $\boldsymbol{x}^*$ being stationary in the problem in a). Establish whether or not the KKT conditions are necessary for a local minimum at $\boldsymbol{x}^*$, and also whether the KKT conditions are sufficient for a feasible vector $\boldsymbol{x}^*$ satisfying the KKT conditions to be a global minimum of the same problem.

(1p) c) Establish whether or not the vector $\boldsymbol{x} = (1, 1/2)^{\mathrm{T}}$ is the projection of $\boldsymbol{z}$ onto the set S.

(3p) Question 4

(the gradient projection method)

Consider the optimization problem to

$$
\begin{aligned}
\text{minimize} \quad & x_1^2 + x_1 x_2 + 2x_2^2 - 10x_1 - 4x_2, \\
\text{subject to} \quad & x_1 + x_2 \leq 3, \\
& 0 \leq x_1 \leq 2, \\
& 0 \leq x_2 \leq 2.
\end{aligned}
$$

Recall that the gradient projection algorithm is a generalization of the steepest descent method to problems over convex sets. Given a point $\boldsymbol{x}_k$, the next point is obtained according to $\boldsymbol{x}_{k+1} = \mathrm{Proj}_X(\boldsymbol{x}_k - \alpha_k \nabla f(\boldsymbol{x}_k))$, where X is the convex set over which we minimize, $\alpha_k > 0$ is the step length in iteration k, and $\mathrm{Proj}_X(\boldsymbol{y}) = \arg\min_{\boldsymbol{x} \in X} \|\boldsymbol{x} - \boldsymbol{y}\|$ denotes the closest point in X to $\boldsymbol{y}$.

Start at the point $\boldsymbol{x}_0 = (2, 1)^{\mathrm{T}}$ and perform two iterations of the gradient projection algorithm with step lengths $\alpha_0 = \frac{1}{2}$ and $\alpha_1 = \frac{1}{4}$. Note that the special form of the feasible region X makes the projection very easy! Is the point obtained a local/global optimum?

(3p) Question 5

(modelling)

You are assigned a number of tiles, each containing a letter from the alphabet. For any letter α your inventory contains N_α (a nonnegative integer) tiles with the letter α. Use the tiles to build words from the collection $w_1, w_2, ..., w_n$. Let $o_{i\alpha}$ be the number of occurrences of letter α in word w_i, $i = 1, \dots, n$. You receive $p_i \geq 0$ points for making word w_i and an additional bonus $b_{ij} \geq 0$ points for making both words w_i and w_j $(i, j = 1, \dots, n)$. Formulate a linear integer program to determine your optimal choice of words. You may construct any word at most once, and use any tile at most once.

Question 6

(true or false)

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

(1p) a) *Claim:* If a function $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex and differentiable then the problem to minimize f over $\mathbb{R}^n$ has a unique optimal solution.

(1p) b) *Claim:* If a function $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex and twice differentiable then its Hessian is positive definite everywhere.

(1p) c) *Claim:* If the function $f : \mathbb{R}^n \to \mathbb{R}$ is concave on $\mathbb{R}^n$ and $c \in \mathbb{R}$, then the set $\{\, \boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) \geq c \,\}$ is convex.

(3p) Question 7

(the Separation Theorem)

The Separation Theorem can be stated as follows.

Suppose that the set $S \subseteq \mathbb{R}^n$ *is closed and convex, and that the point* $\boldsymbol{y}$ *does not lie in* S*. Then, there exist a vector* $\boldsymbol{\pi} \neq \mathbf{0}^n$ *and* $\alpha \in \mathbb{R}$ *such that* $\boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{y} > \alpha$ *and* $\boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{x} \leq \alpha$ *for all* $\boldsymbol{x} \in S$*.*

Establish the theorem using basic results from the course. If you rely on other results when performing your proof of the above theorem, then those results must be stated; they may however be utilized without proof.

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 15–01–13

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:

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

An obvious starting basis is (s_1, x_1) and we can thus begin directly with *Phase II*. Calculating the reduced costs we obtain $\tilde{\mathbf{c}}_N = (0, -1, 1)^{\mathrm{T}}$, meaning that x_3 enters the basis. From the minimum ratio test, we get that x_1 leaves the basis.

Updating the basis we now have (s_1, x_3) in the basis. Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})^{\mathrm{T}}$. meaning that x_2 enters 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 $\tilde{\mathbf{c}}_N = (\frac{2}{5}, \frac{1}{5}, \frac{3}{5})^{\mathrm{T}} \geq 0$, meaning that the current basis is optimal. The optimal solution is thus

$$(x_1, x_2, x_3, s_1, s_2)^{\mathrm{T}} = (0, 1, 3, 0, 0)^{\mathrm{T}},$$

which in the original variables means $(x_1, x_2, x_3)^{\mathrm{T}} = (0, 1, 3)^{\mathrm{T}}$ with optimal objective value $f^\star = 2$.

(1p) b) Calculating the reduced costs of the problem for the optimal basis of the problem from a), we obtain $\tilde{\mathbf{c}}_N = (\frac{3}{5} + \frac{1}{5}\alpha, -\frac{1}{5} - \frac{2}{5}\alpha, \frac{2}{5} - \frac{1}{5}\alpha)^{\mathrm{T}} \geq 0$ meaning that the the optimal solution from a) remains optimal for $-3 \leq \alpha \leq -\frac{1}{2}$.

(3p)
Question 2

(convexity)

This is Theorem 4.3.

Question 3

(KKT optimality conditions)

(1p) a) With $\boldsymbol{z} = (2, 3/2)^{\mathrm{T}}$, the optimization problem to solve is that to

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{z}\|^2, \\ \text{subject to} \quad & x_1 + x_2 \leq 3/2, \\ & x_j \geq 0, \; j = 1, 2. \end{aligned}$$

The objective function is clearly a convex function and the feasible set is a convex set. Hence, the optimization problem is a convex optimization problem.

(1p) b) The KKT conditions for a feasible vector $\boldsymbol{x}^*$ are as follows:

$$\begin{aligned} \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} - \begin{pmatrix} 2 \\ 3/2 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \mu_1(x_1^* + x_2^* - 3/2) &= 0, \\ \mu_2 x_1^* &= 0, \\ \mu_3 x_3^* &= 0, \\ \mu_j &\geq 0, \quad j = 1, 2, 3. \end{aligned}$$

All constraints are affine, so the KKT conditions are necessary for optimality. Since it is a convex optimization problem, the KKT conditions are also sufficient for optimality.

(1p) c) At $\boldsymbol{x}^* = (1, 1/2)^{\mathrm{T}}$, it must hold that $\mu_2 = 0$ and $\mu_3 = 0$. The remaining part of the KKT conditions is then:

$$\mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which has solution $\mu_1 = 1 \geq 0$. Hence, the point $\boldsymbol{x}^*$ is a KKT. From b) we know that it is also an optimal solution.

Question 4

(the gradient projection algorithm)

We have that $\nabla f(\boldsymbol{x}) = (2x_1 + x_2 - 10, x_1 + 4x_2 - 4)^{\mathrm{T}}$. So $\nabla f(\boldsymbol{x}_0) = (-5, 2)^{\mathrm{T}}$ and $\boldsymbol{x}_0 - \alpha_0 \nabla f(\boldsymbol{x}_0) = (9/2, 0)^{\mathrm{T}}$. Performing the projection we get that $\boldsymbol{x}_1 = \mathrm{Proj}_X(\boldsymbol{x}_0 - \alpha_0 \nabla f(\boldsymbol{x}_0)) = \mathrm{Proj}_X\left((9/2, 0)^{\mathrm{T}}\right) = (2, 0)^{\mathrm{T}}$.

It holds $\nabla f(\boldsymbol{x}_1) = (-6, -2)^{\mathrm{T}}$ and $\boldsymbol{x}_1 - \alpha_1 \nabla f(\boldsymbol{x}_1) = (7/2, 1/2)^{\mathrm{T}}$. Performing the projection we get that $\boldsymbol{x}_2 = \mathrm{Proj}_X(\boldsymbol{x}_1 - \alpha_1 \nabla f(\boldsymbol{x}_1)) = \mathrm{Proj}_X\left((7/2, 1/2)^{\mathrm{T}}\right) = (2, 1/2)^{\mathrm{T}}$.

The point $\boldsymbol{x}_2$ is actually a global minimum. This can be verified by either taking another step with the algorithm or by noting that the point is a KKT-point.

(3p) Question 5

(modelling)

For each word w_i, we introduce a binary decision variable x_i such that $x_i = 1$ if and only if word w_i is built. For each pair of words w_i and w_j with $j > i$, a binary variable y_{ij} is used. If $x_i = 0$ or $x_j = 0$ we require that $y_{ij} = 0$. A model can then be written as

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{n} p_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{ij} y_{ij} \\
\text{subject to} \quad & \sum_{i=1}^{n} o_{i\alpha} x_i \le N_\alpha, && \forall \alpha \\
& y_{ij} \le x_i, && i, j = 1, \dots, n, j > i \\
& y_{ij} \le x_j, && i, j = 1, \dots, n, j > i \\
& y_{ij} \in \{0, 1\}, && i, j = 1, \dots, n, j > i \\
& x_i \in \{0, 1\}, && i = 1, \dots, n.
\end{aligned}$$

The program is linear with binary variables.

Question 6

(true or false)

(1p) a) False. Take $f(x) := e^x$.

(1p) b) False. Take $f(x) := x^4$ at $x = 0$.

(1p) c) True. See Proposition 3.65.

(3p) **Question 7**

(the Separation Theorem)

This is Theorem 4.28.
