Chalmers/GU
Mathematics

EXAM

TMA946/MAN280
APPLIED OPTIMIZATION

Date:	03–08–25
Time:	House V, morning
Aids:	Text memory-less calculator
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:	Erik Broman (0740-459022)
Result announced:	03–09–08
	Short answers are also given at the end of
	the exam on the notice board for optimization
	in the MD building.

Exam instructions

When you answer the questions

State your methodology carefully.
Use generally valid methods and theory.

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 linear programming problem to

$$\begin{array}{llll} \text{minimize } z = & 3x_1 + cx_2 + x_3 & & \\ \text{subject to} & 2x_1 \quad\quad + x_3 \geq 3, & & \\ & 2x_1 + 2x_2 + x_3 = 5, & & \\ & x_1, \quad x_2, \quad x_3 \geq 0. & & \end{array}$$

(2p) a) For $c = 1$, solve the problem by using the Simplex method using both Phase I and Phase II.

(1p) b) Find the interval for the values of the constant c such that the optimal basis from problem a) is optimal.

Hint: The interval should include the value $c = 1$.

Some matrix inverses that might come in handy for the problem in a) are

$$\begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0.5 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 2 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 0.5 & 0 \\ -0.5 & 0.5 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -0.5 & 0.5 \end{pmatrix},$$

$$\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.5 & 0 \\ -1 & 1 \end{pmatrix}.$$

(3p)

Question 2

(linear programming theory)

Let A be a matrix of order $m \times n$ and c and b vectors of order $n \times 1$ and $m \times 1$, respectively. Prove that the following LP-problem is *either* infeasible *or* has an optimal objective value of zero.

$$\begin{array}{lll} \text{minimize} & z = c^T x & -b^T y \\ \text{subject to} & Ax & \geq b, \\ & & -A^T y \geq -c, \\ & x, & y \geq 0. \end{array}$$

(3p) Question 3

(modelling)

The company Kvartersbutiken runs a chain of small convenience stores with generous opening hours. They have now decided to establish themselves in Göteborg. In order to know where to place their stores, they have made an extensive market survey. In this survey, the population of Göteborg has been divided into m customer areas, with c_i potential customers in each area. The company has surveyed n possible store locations in their research, and the maximum customer capacity of a store in location j is given by s_j. A store is said to belong to a customer area's *primary* region if the store and the area are really close, and customers will always prefer stores in the primary area. The set P_i lists the stores in the primary region for customer area i. If there is no store in the primary area, or if all these stores are full, then some customers may choose to walk a bit if they really need to buy an item, whereas others will simply not shop. The company assumes that 50% of the potential customers who cannot be served within the primary region will go to a store within the secondary region, while the other 50% will go home without shopping. The stores in the secondary area of region i is given by the set S_i. The annual cost of running a store in location j is r_j, and each customer served will give an annual income of q.

Figure 1 illustrates the different customer areas.

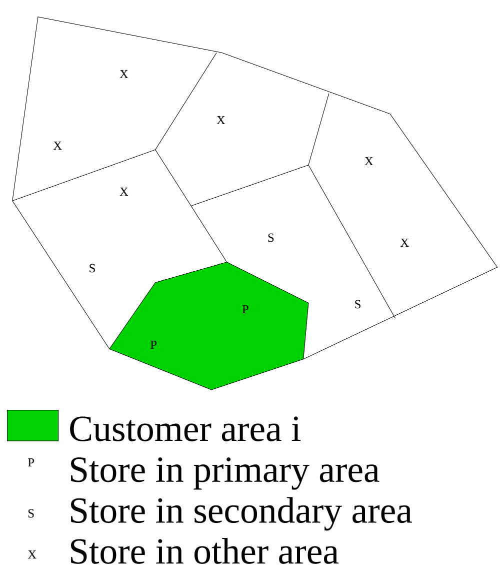

Figure 1: Customer areas.

Your task is to formulate an optimization problem to find where the stores should be located, given the customer behaviour described.

Important note: In order to be able to solve the problem using standard software, the formulated problem should be a *mixed integer linear problem*. In other words, all constraints and the objective function should be linear functions. Specifically, constraints of the type $x = \max\{y, 0\}$ are *not* allowed (although, if this type of constraint is needed, it may be rewritten into the allowed form). Furthermore, conditions written as "x is defined to be 0 if y is 0" is not allowed; this type of logical condition must be rewritten as linear constraints.

Question 4

(definitions)

(2p) a) For a function $f : \Re^n \mapsto \Re$, define (in mathematical notation!) the property of *convexity*.

Define also (in mathematical notation!) the property of *strict convexity* for the function f.

Provide an example function which is convex on $\Re^n$ but *not* strictly convex on $\Re^n$.

Provide, finally, an example function which is strictly convex and differentiable on $\Re^n$, but for which it is *not* true that the Hessian matrix is positive definite everywhere.

(1p) b) Define the term *degenerate basic feasible solution* in linear programming. Introduce all necessary notation.

Question 5

(sensitivity analysis and parametric optimization)

Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(x) := \sum_{j=1}^{n} c_j x_j^2, \\ \text{subject to} \quad & \sum_{j=1}^{n} x_j = b, \end{aligned}$$

where c_j ($j = 1, \ldots, n$) and b all are positive constants.

(2p) a) Establish that a unique optimal solution exists to the above problem, and describe the optimal solution as an *explicit* function of the constants c_j ($j = 1, \ldots, n$) and b.

Hint: Utilize the KKT conditions for the problem. End up with an explicit formula of the form $x_j^* = p_j(c, b)$ for some function p_j.

Interpret the expression. Answer in particular the following question: For each $j = 1, \ldots, n$, is $x_j^* = p_j(c, b)$ an increasing/decreasing function of c_j and/or b? Why is it natural?

(1p) b) We now let b be a parameter, that is, we allow the right-hand side of the constraint to vary. (The constants c_j remain constant, and we write $x_j^* = p_j(b)$ to reflect this.) We are interested in learning about the sensitivity of the optimal solution as a function of b, and will have good use for the explicit formula derived in (a).

Establish that the optimal solution x^* is differentiable at b. Determine the expression of the derivative of $x_j^* = p_j(b)$ as a function of b, that is, determine the expression of $\nabla_b x_j^* = p_j'(b)$.

Interpret the expression of the above derivative. In particular, answer the following questions: (i) What is the sign of the derivative? (ii) Which variable x_j increases/decreases the fastest when the value of b increases? Why?

(3p) Question 6

(steepest descent method)

Consider the following (rather simple!) optimization problem:

$$\text{minimize } \frac{1}{2}\|x\|^2 := \frac{1}{2}\sum_{i=1}^{n} x_i^2. \qquad (1)$$

We would like to apply the steepest descent algorithm with a fixed stepsize (that is, the line search is replaced by a fixed step in each iteration) to solve this problem:

initialization: choose a starting point $x^0 \in \Re^n$, and a constant $s > 0$.

main step: $x^{t+1} := x^t - s\nabla f(x^t)$; let $t := t + 1$ and repeat until a convergence criterion is met.

[In the description of the algorithm, f is the objective function, $\|x\|^2$ in our case.]

(2p) a) Find the set of all pairs (x^0, s) in $\Re^n \times \Re$, for which the algorithm described converges towards the globally optimal solution to the problem (1).

(1p) b) It is known that the steepest descent algorithm is guaranteed to converge to stationary points of the continuously differentiable function $f : \Re^n \mapsto \Re$ from *any* starting point x^0 if $0 < s < 2/L$, where L is a *Lipschitz constant* for ∇f on $\Re^n$. (We note that ∇f is Lipschitz continuous on $\Re^n$ if there exists a constant $L \geq 0$ such that for every pair $x, y \in \Re^n$ it holds that

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|.)$$

Is the convergence set given by this inequality bigger, smaller, or equal to the one you have obtained in a)?

Hint: What is the Lipschitz constant L for the objective function in (1)?

Question 7

(convexity)

For any set $S \subset \Re^n$ the *convex hull* of S, denote $\text{hull}\,(S)$, is the set of all possible convex combinations of points in S. An illustration of the convex hull is given in Figure 2.

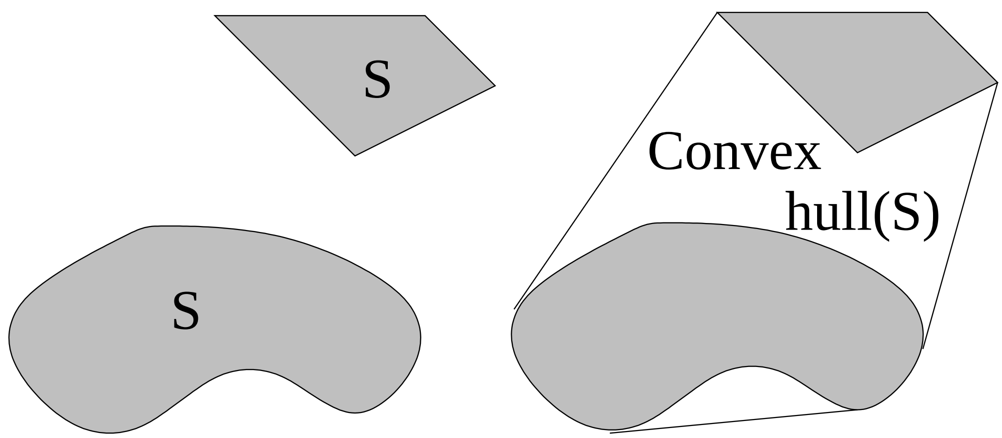

Figure 2: A set S consisting of two disjoint sets (left) and its convex hull(right).

(2p) a) Let S_1 and S_2 be arbitrary subsets of $\Re^n$. Show that $\text{hull}\,(S_1 \cap S_2) \subseteq \text{hull}\,(S_1) \cap \text{hull}\,(S_2)$.

(1p) b) Prove by means of a counterexample that $\text{hull}\,(S_1 \cap S_2) = \text{hull}\,(S_1) \cap \text{hull}\,(S_2)$ does *not* hold in general.

Good luck!

TMA 946 030825

Solution to question 1: (a) By introducing a slack variable x_4 and two artificial variables x_5 and x_6 we get the Phase I problem

$$\begin{array}{llllllll} \text{minimize } w = & & & & & +x_5 & +x_6 & \\ \text{subject to} & 2x_1 & & +x_3 & -x_4 & +x_5 & & = 3, \\ & 2x_1 & +2x_2 & +x_3 & & & +x_6 & = 5, \\ & x_1, & x_2, & x_3, & x_4, & x_5, & x_6 & \geq 0. \end{array}$$

Let $x_B = [x_5, x_6]$ and $x_N = [x_1, x_2, x_3, x_4]$ be the initial basic and nonbasic vector respectively. The reduced costs of the nonbasic variables then become

$$c_N^T - c_B^T B^{-1} N = [-4, -2, -2, 1],$$

and thus x_1 is the entering variable. Further, we have

$$B^{-1}b = [3, 5]^T,$$
$$B^{-1}N_1 = [2, 2]^T,$$

which gives

$$\operatorname*{arg\,min}_{j, (B^{-1}N_1)_j > 0} \frac{(B^{-1}b)_j}{(B^{-1}N_1)_j} = 1,$$

so x_5 is the leaving variable. The new basic and nonbasic vectors are $x_B = [x_1, x_6]$ and $x_N = [x_5, x_2, x_3, x_4]$, and the reduced costs of the nonbasic variables become

$$c_N^T - c_B^T B^{-1} N = [2, -2, 0, -1],$$

so x_2 is the entering variable, and

$$B^{-1}b = [1.5, 2]^T,$$
$$B^{-1}N_2 = [0, 2]^T,$$

which gives

$$\operatorname*{arg\,min}_{j, (B^{-1}N_2)_j > 0} \frac{(B^{-1}b)_j}{(B^{-1}N_2)_j} = 2,$$

and thus x_6 is the leaving variable. The new basic and nonbasic vectors become $x_B = [x_1, x_2]$ and $x_N = [x_5, x_6, x_3, x_4]$, and the reduced costs of the nonbasic variables are

$$c_N^T - c_B^T B^{-1} N = [1, 1, 0, 0],$$

so $x_B = [x_1, x_2]$ is an optimal basis to the Phase I problem, and $w^* = 0$. This means that $x_B = [x_1, x_2]$ gives a basic feasible solution to the Phase II problem, that is,

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

If $x_B = [x_1, x_2]$ and $x_N = [x_3, x_4]$ we get the reduced costs

$$c_N^T - c_B^T B^{-1} N = [-0.5, 1],$$

which means that x_3 is the entering variable, and

$$B^{-1}b = [1.5, 1]^T,$$
$$B^{-1}N_1 = [0.5, 0]^T,$$

which gives

$$\underset{j,\,(B^{-1}N_1)_j>0}{\arg\min} \frac{(B^{-1}b)_j}{(B^{-1}N_1)_j} = 1,$$

so x_1 is the leaving variable. We get $x_B = [x_3, x_2]$ and $x_N = [x_1, x_4]$, and the reduced costs become

$$c_N^T - c_B^T B^{-1} N = [1, 0.5],$$

so $x_B = [x_3, x_2]$ is an optimal basis, and since

$$B^{-1}b = [3, 1]^T$$

an optimal solution is given by

$$x^* = [x_1, x_2, x_3] = [0, 1, 3],$$

and $z^* = 4$.

(b) The reduced cost is

$$c_N^T - c_B^T B^{-1} N = [3, 0] - [1, c] \begin{pmatrix} 1 & 0 \\ -0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix} = [1, 1 - 0.5c]$$

which gives $c \leq 2$. □

Solution to problem 2: The dual to the given LP-problem is

$$\begin{array}{llll} \text{maximize } w = & -cx & +b^T y & \\ \text{subject to} & Ax & & \geq b, \\ & & -A^T y & \geq -c^T, \\ & x, & y & \geq 0, \end{array}$$

which means that the primal problem and the dual problem has the same set of feasible solutions. Thus, the primal problem is feasible if and only if the dual problem is feasible and weak duality gives that the problem cannot be unbounded. Suppose that $z^* > 0$. Strong duality gives that there exists a dual solution such that $w^* = z^*$. But this dual solution is feasible to the primal problem and gives lower value of z than z^*, which contradicts the optimality of z^*. Now suppose that $z^* < 0$. The primal optimal solution is feasible to the dual problem and gives a dual objective value that is higher than z^*, which contradicts weak duality. □

3 solution

In order to formulate the problem, we introduce the following notation:
Parameters:

- m Number of customer areas.
- n Number of possible store locations.
- P_i Set of locations in area i's primary region
- S_i Set of locations in area i's secondary region
- c_i Potential customers in area i
- s_j Maximum capacity of store j.
- r_j Annual cost of ruining a store at location j.
- q Annual income per customer

Variables:

- y_j Binary variable indicating if a store is opened at location i
- x_{ij} Customers from area i shopping at location j.

$$\max \quad \sum_{i=1}^{m} \sum_{j \in P_i \cup S_i} x_{ij} q - \sum_{j=1}^{n} r_j y_j \tag{1.1}$$

$$\text{s.t} \quad \sum_{j \in P_i} x_{ij} + \sum_{j \in S_i} 2x_{ij} \leq c_j, \quad i = 1, \ldots, M \tag{1.2}$$

$$\sum_{i=1}^{m} x_{ij} \leq y_j s_j, \quad j = 1, \ldots, n \tag{1.3}$$

Here, (1.1) measures the income from served customers reduced with the running costs. Equation (1.2) guarantees that more customers does not come from an area than what is actually there. Equation (1.3) makes sure that people only shop in open stores.

4 a) A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex on $\mathbb{R}^n$ if for every $x, y \in \mathbb{R}^n$ and $\lambda \in (0,1)$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

A function $f: \mathbb{R}^n \to \mathbb{R}$ is strictly convex on $\mathbb{R}^n$ if for every $x, y \in \mathbb{R}^n$ with $x \neq y$ and $\lambda \in (0,1)$,

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y).$$

Every affine function, $f(x) = c^T x - q$, $c \in \mathbb{R}^n$, $q \in \mathbb{R}$, is convex on $\mathbb{R}^n$, but <u>not</u> strictly convex.

The function $f(x) = x^4$ is strictly convex and differentiable on $\mathbb{R}$, but since $f''(0) = 0$, its Hessian (here, second derivative) is not positive everywhere.

b) Given the system $Ax = b$; $x \geq 0^n$ where $A \sim m \times n$, $b \in \mathbb{R}^m$, $n > m$ and A has full row rank, a basic feasible solution is a solution to $Bx_B = b$, where $x_B \geq 0^m$, B being an invertible square $m \times m$ matrix formed as follows: $A = (B, N)$; $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$, that is, as a collection of m column vectors of A. x_B is degenerate if x_B contains a zero element.

5. a) With λ being the Lagrangemultiplier, the KKT conditions are:

$$2c_j x_j - \lambda = 0 \quad , \quad j = 1, \dots, n \qquad (1)$$

$$\sum_j x_j = b \qquad (2)$$

From (1) follows that $x_j = \frac{\lambda}{2c_j}$, $j = 1, \dots, n$,

so by (2), $\frac{\lambda}{2} \sum_j \frac{1}{c_j} = b$, and so

$\lambda^* = \frac{2b}{\sum_j 1/c_j}$ must hold. Hence,

$x_j^* = b \Big/ \left(c_j \cdot \sum_k 1/c_k\right)$ is a KKT point. It is easy to check that f is strictly convex and that it is a convex problem, so x^* is unique.

$\therefore$ $x_j^* = p_j(c, b)$, where

$$p_j(c, b) = \frac{b}{c_j \cdot \sum_k \frac{1}{c_k}}$$

p_j is increasing in b. (Natural, as b can be considered as a demand that must be fulfilled, at least cost.)

p_j is decreasing in c_j. (Natural, because it makes x_j less favorable compared to the other variables.)

b) Consider the function

$$p_j(c,b) = \frac{b}{c_j \cdot \sum_k \frac{1}{c_k}}, \quad j=1,\dots,n,$$

over $b > 0$, $c > 0^n$. With $c > 0^n$ fixed, it is clearly a linear function of b. It is therefore differentiable, with

$$\frac{\partial p_j(c,b)}{\partial b} = \frac{1}{c_j \cdot \sum_k \frac{1}{c_k}}.$$

$$\therefore \quad \nabla_b x_j^* = p_j'(b) = \frac{\partial p_j(c,b)}{\partial b} = \frac{1}{c_j \cdot \sum_k \frac{1}{c_k}}.$$

The derivative is positive, showing (as discussed in a)) that the variables x_j <u>all</u> increase at the optimum solution when the value of b increases.

Question 6.

a) $\nabla\left(\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) = X$, therefore

$$X^{t+1} = X^t - sX^t = (1-s)^{t+1}X^0$$

The latter converges either

$\forall s \geq 0$, if $X^0 = 0$ [trivial case]

or

$\forall x^0 \in \mathbb{R}^n$, if $|1-s| < 1$, i.e. $0 < s < 2$.

b) Trivially, $\|\nabla f(x) - \nabla f(y)\| \leq 1 \cdot \|x-y\|$, i.e. $L = 1$ (Lipschitz constant),

Therefore, from the general theory it follows that algorithm converges

$\forall\ 0 < s < 2/1 = 2$,

which is the same as we obtained in a) [excluding trivial case].

Question 7.

a) Let $x \in hull(S_1 \cap S_2)$, i.e.
$x = \lambda y + (1-\lambda) z$ for some $0 \le \lambda \le 1$,
$y \in S_1 \cap S_2$, $z \in S_1 \cap S_2$.

Then, $x \in hull(S_1)$ (since $y, z \in S_1$)
& $x \in hull(S_2)$ (since $y, z \in S_2$),
finishing the proof.

b) Let $S_1 = \{-1, 1\} \subseteq \mathbb{R}$
$S_2 = \{0\} \subset \mathbb{R}$

Then, $hull(S_1) = [-1, 1]$
$hull(S_2) = \{0\}$ so that
$hull(S_1) \cap hull(S_2) = [-1, 1]$, while
$hull(S_1 \cap S_2) = hull(\emptyset) = \emptyset$.