Chalmers/GU
Mathematics

EXAM

TMA947/MAN280
APPLIED OPTIMIZATION

Date:	04–06–02
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:	Alex Herbertsson (0740-459022)
Result announced:	04–06–11 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 program

$$\begin{array}{llrrrl} \text{minimize} & z = & x_1 & +2x_2 & +3x_3 & \\ \text{subject to} & & 2x_1 & -5x_2 & +x_3 & \geq 2, \\ & & 2x_1 & -x_2 & +2x_3 & \leq 4, \\ & & x_1, & x_2, & x_3 & \geq 0. \end{array}$$

(2p) a) Solve the problem by using Phase I & II of the Simplex method.

(1p) b) Is the solution obtained unique? (Motivate!)

In order to calculate necessary matrix inverses the following basic identity might be useful:

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

(3p) ## Question 2

(modelling)

Figure 1 describes two production processes by which we can make the product D from the raw materials A, B, and C.

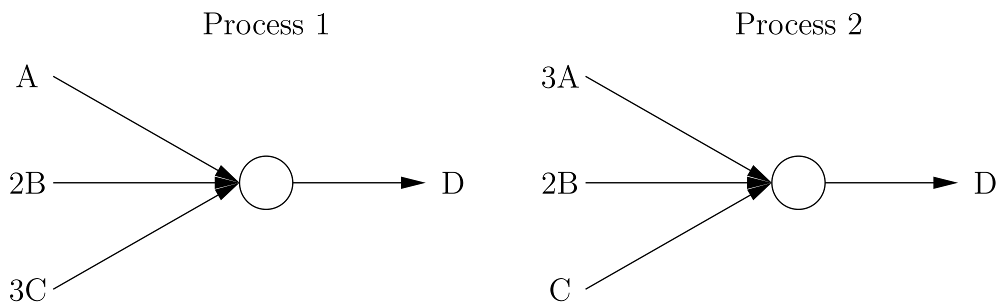

Figure 1: The production processes.

The figure illustrates that process 1 requires 1 unit of raw material A, 2 units of raw material B, and 3 units of raw material C in order to produce 1 unit of product D, and process 2 requires 3 units of A, 2 units of B, and 1 unit of C in order to produce one unit of D. The capacities of the processes 1 and 2 are p_1 and p_2, respectively, of units of the product D. Further the processes 1 and 2 are associated with the fixed start-up costs f_1 and f_2. For example, if process 1 is used for production, then f_1 must be paid independently of how many units of D that are actually produced. The cost per unit of the raw materials A, B, and C are c_A, c_B, and c_C, respectively. There are three demand centers that require d_1, d_2, and d_3 units, respectively, of product D. We may assume that the transportation cost is negligible compared to all other costs.

Formulate a *linear integer programming model* (that is, if the integer requirements are relaxed we shall end up with an ordinary linear program) for finding the production quantities that minimize the total cost given that demand is fulfilled.

Question 3

(Newton's algorithm)

An engineer has decided to verify numerically that the exponential function $x \mapsto \exp(x) = e^x$ grows faster than any polynomial. In order to do so he/she studies the optimization problem to

$$\text{minimize} \;\; f(x) = x^\alpha - \exp(x), \tag{1}$$

where α is the highest power of the polynomial (we assume it is an even, positive integer number). The engineer uses a Newton method (with unit steps!) to solve the problem. He/she argues that if the exponential function grows faster than any polynomial, then the sequence $\{x_k\}$ generated by the method should converge to infinity, because the objective function f can be decreased indefinitely by increasing the value of x.

(1p) a) State the Newton iteration explicitly for the given problem (1).

(1p) b) Find the error in the engineer's reasoning and formally explain it.

(1p) c) Construct a numerical example (that is, choose a value of $\alpha \in \{2, 4, \dots\}$, and a starting point of the Newton algorithm) illustrating the engineer's error in reasoning.

Question 4

(claims about optimality)

Each of the three questions below are to be answered independently. For each of them we now describe the task to be performed: After the problem description, a claim is made. Given the properties stated before the claim, the claim is *not* true. However, under additional properties of the problem the claim *is* true. Your task is to describe a *reasonable and mild* additional set of properties that makes the claim valid. (All claims are based on basic results in the course notes.) In addition to providing these additional properties, you must also state *why* this property is needed, by providing a counter-example to the claim for the case when the property is not present.

(1p) a) Consider the standard LP problem to

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}^n, \end{aligned}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n$, and $\boldsymbol{b} \in \mathbb{R}^m$. Suppose further that there exist feasible solutions to this problem.

Claim: There exists at least one optimal solution to this problem.

(1p) b) Consider the problem to

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

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is continuously differentiable and lower bounded on $\mathbb{R}^n$. Suppose that to this problem we apply the steepest descent algorithm with the Armijo step length rule, starting at some $\boldsymbol{x}_0 \in \mathbb{R}^n$ and generating a sequence $\{\boldsymbol{x}_k\}$ of iterates.

Claim: The following is true:

- $\{f(\boldsymbol{x}_k)\} \downarrow \bar{f} \in \mathbb{R}$;
- $\{\nabla f(\boldsymbol{x}_k)\} \to \boldsymbol{0}^n$;
- $\{\boldsymbol{x}_k\}$ has at least one accumulation point, $\bar{\boldsymbol{x}} \in \mathbb{R}^n$; for each accumulation point $\bar{\boldsymbol{x}}$, $f(\bar{\boldsymbol{x}}) = \bar{f}$ and $\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0}^n$ holds.

(1p) c) Consider the problem to

$$\begin{aligned} &\text{minimize} \;\; f(\boldsymbol{x}), \\ &\text{subject to} \;\; g(\boldsymbol{x}) \leq b, \end{aligned}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $g : \mathbb{R}^n \mapsto \mathbb{R}$ are continuous functions, and $b \in \mathbb{R}$. Suppose that this problem has a globally optimal solution, $\boldsymbol{x}^*$, and that $g(\boldsymbol{x}^*) < b$ holds.

Claim: The vector $\boldsymbol{x}^*$ is also a globally optimal solution to the unconstrained problem to

$$\begin{aligned} &\text{minimize} \;\; f(\boldsymbol{x}), \\ &\text{subject to} \;\; \boldsymbol{x} \in \mathbb{R}^n. \end{aligned}$$

Question 5

(duality)

Consider the optimization problem to

$$\begin{aligned} \text{minimize} \;\; & f(x, y) = y, \\ \text{s.t.} \;\; & (x-1)^2 + y^2 \leq 1, \\ & (x+1)^2 + y^2 \leq 1, \end{aligned} \tag{1}$$

where $x, y \in \mathbb{R}$.

(1p) a) Find *every* point of global and local minimum (you may do this graphically). Is this a convex problem? Does it verify Slater's CQ or the linear independence CQ (LICQ)?

(1p) b) Derive the expression of the Lagrangian dual function $q : \mathbb{R}^2_+ \mapsto \mathbb{R} \cup \{-\infty, +\infty\}$ associated with the Lagrangian relaxation of both constraints of the problem (1).

(1p) c) Show that strong duality holds, that is, that $f^* = \sup_{\lambda \in \mathbb{R}^2_+} q(\lambda)$ holds, where f^* is the optimal value of the primal problem (1).

Question 6

(convexity)

Throughout the course we have stressed that convexity is a crucial property of functions when analyzing optimization models in general and studying optimality conditions in particular. There are, however, certain properties of convex functions that are shared also by classes of non-convex functions. The purpose of this question is to relate the convex functions to two such classes of non-convex functions by means of some example properties.

Suppose that $S \subseteq \mathbb{R}^n$ and that $f : \mathbb{R}^n \mapsto \mathbb{R}$ is continuous on S.

(1p) a) Suppose further that f is continuously differentiable on S (C^1 on S). We say that the function f is *pseudo-convex* on S if, for every $\boldsymbol{x}, \boldsymbol{y} \in S$,

$$\nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}) \geq 0 \qquad \Longrightarrow \qquad f(\boldsymbol{y}) \geq f(\boldsymbol{x}).$$

Establish the following two statements: (1) every differentiable, convex function on $\mathbb{R}^n$ is pseudo-convex on $\mathbb{R}^n$ (that is, "convexity implies pseudo-convexity"); (2) the reverse statement ("pseudo-convexity implies convexity") is not true. *Hint:* On the statement (2) you may construct an explicit or graphical counter-example.

(1p) b) A well-known property of a differentiable convex function is its role in necessary and sufficient conditions for globally optimal solutions. Suppose now that S is convex. If f is a convex function on $\mathbb{R}^n$ which is in C^1 on S then the following statement holds (Theorem 4.21 in the course notes):

$$\boldsymbol{x}^* \text{ is a global minimum of } f \text{ over } S \iff \nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}^*) \geq 0, \forall \boldsymbol{y} \in S.$$

Establish that this equivalence relation still holds if the convexity of f is replaced by the pseudo-convexity of f.

(1p) c) Let S be convex. We say that the function f is *quasi-convex* on S if its level sets are convex. In other words, f is *quasi-convex* on S if

$$\operatorname{lev}_f^S(b) := \{\, \boldsymbol{x} \in S \mid f(\boldsymbol{x}) \leq b \,\}$$

is convex for every $b \in \mathbb{R}$.

Establish the following two statements: (1) every convex function on S is quasi-convex on S (that is, "convexity implies quasi-convexity"); (2) the reverse statement ("quasi-convexity implies convexity") is not true. *Hint:* On the statement (2) you may construct an explicit or graphical counter-example.

(3p) Question 7

(Lagrangian duality)

Consider the optimization problem to find

$$\begin{aligned} f^* := \underset{x}{\text{infimum}} \;\; & f(\boldsymbol{x}), \\ \text{subject to} \;\; & g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, \dots, m, \\ & \boldsymbol{x} \in X, \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $g_i : \mathbb{R}^n \mapsto \mathbb{R}$ $(i = 1, 2, \dots, m)$ are given functions, $X \subseteq \mathbb{R}^n$; we assume that $-\infty < f^* < \infty$. For an arbitrary vector $\boldsymbol{\mu} \in \mathbb{R}^m$, we define the *Lagrange function*

$$L(\boldsymbol{x}, \boldsymbol{\mu}) := f(\boldsymbol{x}) + \sum_{i=1}^{m} \mu_i g_i(\boldsymbol{x}) = f(\boldsymbol{x}) + \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}).$$

We call the vector $\boldsymbol{\mu}^* \in \mathbb{R}^m$ a *Lagrange multiplier* if it is non-negative and if $f^* = \inf_{x \in X} L(\boldsymbol{x}, \boldsymbol{\mu}^*)$ holds.

For the problem (1), establish the following theorem on global optimality conditions in the absence of a duality gap:

The vector $(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ is a pair of optimal primal solution and Lagrange multiplier if and only if

$$\boldsymbol{\mu}^* \geq \boldsymbol{0}^m, \qquad \textit{(Dual feasibility)} \tag{2a}$$

$$\boldsymbol{x}^* \in \arg \min_{x \in X} L(\boldsymbol{x}, \boldsymbol{\mu}^*), \qquad \textit{(Lagrangian optimality)} \tag{2b}$$

$$\boldsymbol{x}^* \in X, \quad \boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}^m, \qquad \textit{(Primal feasibility)} \tag{2c}$$

$$\mu_i^* g_i(\boldsymbol{x}^*) = 0, \quad i = 1, \dots, m. \qquad \textit{(Complementary slackness)} \tag{2d}$$

Good luck!

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280 APPLIED OPTIMIZATION

Date: 04–06–02
Examiner: Michael Patriksson

Question 1

(the Simplex method)

a) By introducing a surplus-variable s_1 and a slack-variable s_2 the standard form of the problem is to

$$\begin{array}{lrrrrrl} \text{minimize} & z = x_1 & +2x_2 & +3x_3 & & & \\ \text{subject to} & 2x_1 & -5x_2 & +x_3 & -s_1 & & = 2, \\ & 2x_1 & -x_2 & +2x_3 & & +s_2 & = 4, \\ & x_1, & x_2, & x_3, & s_1, & s_2 & \geq 0. \end{array}$$

By introducing an artificial variable in the first constraint and solving the Phase I problem we get the BFS $\boldsymbol{x}_B = (x_1, s_2)$. This BFS also gives the optimal solution $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}} = (1, 0, 0)^{\mathrm{T}}$ to the original problem.

b) This is the unique optimal solution since the reduced costs of the non-basic variables $\boldsymbol{x}_N = (x_2, x_3, s_1)^{\mathrm{T}}$ are all strictly positive [$\tilde{\boldsymbol{c}}_N = (4.5, 2.5, 0.5)^{\mathrm{T}}$].

Question 2

(modelling)

Introduce the following variables:

$$x_{ij} = \text{the number of D from process } i = 1, 2 \text{ sent to demand center } j = 1, 2, 3,$$

$$z_i = \begin{cases} 1, & \text{if process } i = 1, 2 \text{ is used,} \\ 0, & \text{otherwise.} \end{cases}$$

The linear integer programming problem is then to

$$\begin{array}{ll} \text{minimize} & (c_A + 2c_B + 3c_C) \displaystyle\sum_{j=1}^{3} x_{1j} + (3c_A + 2c_B + c_C) \sum_{j=1}^{3} x_{2j} + \sum_{i=1}^{2} f_i z_i \\ \text{subject to} & \displaystyle\sum_{i=1}^{2} x_{ij} \geq d_j, \quad j = 1, 2, 3, \\ & \displaystyle\sum_{j=1}^{3} x_{ij} \leq p_i z_i, \quad i = 1, 2, \\ & z_i \in \mathbb{B}, \quad x_{ij} \in \mathbb{Z}_+, \quad i = 1, 2, \quad j = 1, 2, 3. \end{array}$$

Question 3

(Newton's algorithm)

a) Newton's equation:

$$x_{k+1} = x_k - \frac{\alpha x^{\alpha-1} - \exp(x)}{\alpha(\alpha-1)x^{\alpha-2} - \exp(x)}$$

b) The objective function of the problem is not convex in general [may be verified by analysing the sign of the Hessian $\alpha(\alpha-1)x^{\alpha-2} - \exp(x)$]. Since the convergence of the Newton method is local in nature, the method is most likely to converge to the nearest local minimum. The engineer thus wrongly assumes the global convergence of the Newton method.

c) Probably the simplest counter-example is obtained by taking $x_0 = 1$, $\alpha = 2$. These initial values cause the Newton's method to generate an oscillating sequence $x_{2k-1} = 0$, $x_{2k} = 1$, $k = 1, 2, \dots$.

Question 4

(claims about optimality)

a) Additional property: the polyhedron is bounded.

Counter-example: the problem to maximize x_1 subject to $x_1 \geq 0$ has no optimal solution.

b) Additional property: f is weakly coercive.

Counter-example:

$$\text{minimize } f(x) = \begin{cases} -x, & x \leq 1, \\ 1/x, & x \geq 1 \end{cases}$$

is in C^1 on $\mathbb{R}$ and lower bounded. If we apply the steepest descent method on it, however, we obtain $\{x_k\} \to \infty$ while $\{f'(x_k)\} \to 0$.

c) Additional property: f and g are convex functions.

Counter-example: $f(x) = x^5 - 100x^3$; $g(x) = -x$; and $b = -6$ (that is, the constraint is $x \geq 6$).

The below plot shows the appearance of the function f in the interval $[-12, 12]$; clearly, the optimal solution to the constrained problem is $x^* \approx 7.5$, and $g(x^*) < b$ holds, but if we remove the constraint we see from the figure that there is no optimal solution to the problem—we may let $f(x)$ tend to minus infinity by letting x tend to minus infinity.

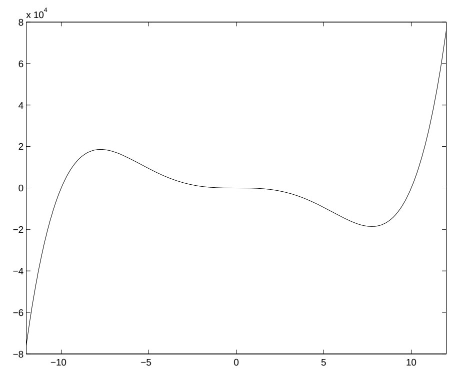

Figure 1: The function $f(x) = x^5 - 100x^3$ on an interval.

Question 5

(duality)

a) The only point that is feasible in the problem is the point $(x, y) = (0, 0)$ (easily verified graphically); thus it is the only locally and globally optimal solution. The problem is convex (both the objective function and the less-than-or-equal-to constraints are convex). It however does not satisfy the Slaters's CQ (there are no strictly feasible points), or LICQ (the gradients of the active constraints at the only feasible point are linearly dependent). An alternative argument is that the locally optimal solution $(0, 0)^{\mathrm{T}}$ is not a KKT point, which means that CQs cannot hold.

b) Introducing the Lagrange multipliers μ and λ for the constraints of the problem, we get

$$q(\lambda,\mu) = \inf_{(x,y)\in\mathbb{R}^2} \left\{ y + \lambda[(x-1)^2 + y^2 - 1] + \mu[(x+1)^2 + y^2 - 1] \right\}. \quad (1)$$

If $\lambda = \mu = 0$ we get $q(\lambda,\mu) = -\infty$; it remains thus to calculate q for assuming $\lambda+\mu > 0$. From the necessary (and sufficient in this convex case) optimality conditions we get:

$$\begin{cases} 2\lambda[x-1] + 2\mu[x-1] = 0 \\ 1 + 2\lambda y + 2\mu y = 0 \end{cases} \equiv \begin{cases} x = \dfrac{\lambda-\mu}{\lambda+\mu} \\ y = -\dfrac{1}{2(\lambda+\mu)}. \end{cases}$$

Substituting this into (1) we finally obtain

$$q(\lambda,\mu) = -\frac{1}{4(\lambda+\mu)} - \frac{(\lambda-\mu)^2}{\lambda+\mu}.$$

c) Show that the strong duality holds, that is, $z^* = \sup_{\lambda\in\mathbb{R}^2_+} q(\lambda)$, where z^* is the optimal value of the primal problem.

Clearly, in our case $z^* = 0$. Thus, by the weak duality, or from the explicit formula for the dual function, we have that for all $(\lambda,\mu) \in \mathbb{R}^2_+$ it holds that $q(\lambda,\mu) < 0$. Still, $\lim_{\lambda\to+\infty} q(\lambda,\lambda) = 0$, which means that $\sup_{(\lambda,\mu)\in\mathbb{R}^2_+} q(\lambda,\mu) = 0 = z^*$.

Question 6

(convexity)

a) (1) We utilize the following characterization of convexity of f on $\mathbb{R}^n$:

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{x}), \qquad \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n.$$

It follows that if $\boldsymbol{x}$ and $\boldsymbol{y}$ are such that $\nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{x}) \geq 0$ then $f(\boldsymbol{y}) \geq f(\boldsymbol{x})$ also holds; hence, f is pseudo-convex on $\mathbb{R}^n$.

(2) The function in the below figure is of a form often referred to as "unimodal," that is, it has a unique minimum and it increases both to the right and left of this minimum. If such a function is differentiable then it is also pseudo-convex (check this!).

It is however not convex.

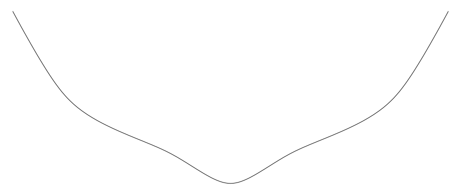

Figure 2: A unimodal function.

b) The result in the direction of $\Longrightarrow$ (the necessary condition) is true for every differentiable function. The result in the direction of $\Longleftarrow$ (the sufficient condition) follows immediately from the definition of pseudo-convexity.

c) (1) This result is Proposition 3.48 in the Course Notes.

(2) A unimodal function has convex level sets, and so the example in the above figure works as a conuter-example here as well.

Question 7

(Lagrangian duality)

The proof of the consistency of the global optimality conditions be found in Theorem 7.6 in the Course Notes.