Chalmers/GU
Mathematics

EXAM

TMA947/MAN280
APPLIED OPTIMIZATION

Date:	06–03–06
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:	Peter Lindroth (0762-721860)
Result announced:	06–03–24 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 methods and theory.
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:

$$
\begin{array}{lrrrl}
\text{minimize} & z = -x_1 & +2x_2 & +x_3, & \\
\text{subject to} & 2x_1 & +x_2 & -x_3 & \leq 7, \\
& -x_1 & +2x_2 & +3x_3 & \geq 3, \\
& x_1, & x_2, & x_3 & \geq 0.
\end{array}
$$

(2p) a) Solve this problem by using phase I and phase II of the simplex method.
[*Aid:* Utilize the identity

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

for producing basis inverses.]

(1p) b) Using the information from your solution to the problem above, state whether there is an optimal solution to the corresponding dual problem or not. Motivate! Do not make any additional calculations!

(3p)

Question 2

(The Karush–Kuhn–Tucker conditions)

Consider the problem to

$$
\begin{aligned}
&\text{minimize } f(\boldsymbol{x}), \\
&\text{subject to } g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, \dots, m,
\end{aligned}
\tag{1}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, are given differentiable functions. Prove the following result.

Assume that at a given point $\boldsymbol{x}^* \in S$ Abadie's constraint qualification holds. If $\boldsymbol{x}^* \in S$ is a local minimum of f over S then there exists a vector $\boldsymbol{\mu} \in \mathbb{R}^m$ such that

$$\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(\boldsymbol{x}^*) = \mathbf{0}^n, \tag{2a}$$

$$\mu_i g_i(\boldsymbol{x}^*) = 0, \qquad i = 1, \dots, m, \tag{2b}$$

$$\boldsymbol{\mu} \geq \mathbf{0}^m. \tag{2c}$$

In other words,

$$\left.\begin{array}{r} \boldsymbol{x}^* \textit{ local minimum of } f \textit{ over } S \\ \textit{Abadie's CQ holds at } \boldsymbol{x}^* \end{array}\right\} \implies \exists \boldsymbol{\mu} \in \mathbb{R}^m : \text{ (2) } \textit{holds.}$$

You may refer (simply by giving its name) to any additional theorems that you wish to utilize in your proof.

Question 3

(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.

(1p) a) The vector $\boldsymbol{p} := (1, -1)^{\mathrm{T}}$ is a descent direction for $f(\boldsymbol{x}) := (x_1 + x_2^2)^2$ at $\boldsymbol{x} = (1, 0)^{\mathrm{T}}$.

(1p) b) The vector $\boldsymbol{x} := (0, 1)^{\mathrm{T}}$ satisfies the KKT conditions for the problem to

$$\begin{aligned} \underset{x \in \mathbb{R}^2}{\text{maximize}} \quad & f(x) := \ln(x_1 + x_2), \\ \text{subject to} \quad & x_1 + 2x_2 \leq 2, \\ & x_2 \geq 0. \end{aligned}$$

(1p) c) Consider the linear program defined by

$$\begin{aligned} f(\boldsymbol{b}) := \underset{x \in \mathbb{R}^n}{\text{maximum}} \quad & \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}, \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \qquad (1) \\ & \boldsymbol{x} \geq \boldsymbol{0}^n, \end{aligned}$$

where $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, and $\boldsymbol{b} \in \mathbb{R}^m$, and let $\boldsymbol{y} \in \mathbb{R}^m$ be the vector of dual variables associated with the constraints (1). Assume that for the given value of $\boldsymbol{b}$ there exists a non-degenerate primal optimal BFS. Hence, there exists a unique optimal dual solution $\boldsymbol{y}^* \in \mathbb{R}^m$. Suppose that $y_1^* > 0$ and let $\overline{\boldsymbol{b}} \in \mathbb{R}^m$ be such that $b_1 < \overline{b}_1 < b_1 + \varepsilon$ and $\overline{b}_i = b_i$, $i = 2, \ldots, m$. For small values of $\varepsilon > 0$ it holds that $f(\overline{\boldsymbol{b}}) > f(\boldsymbol{b})$.

Question 4

(nonlinear programming)

Consider the nonlinear program to

$$\begin{aligned} &\text{minimize } f(\boldsymbol{x}), \\ &\text{subject to } h_j(\boldsymbol{x}) = 0, \qquad j = 1, \dots, \ell, \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, \dots, \ell$, are given differentiable functions.

(1p) a) When is the problem (1) a convex problem?

(2p) b) Suppose that having utilized a nonlinear programming solver the printout is a vector $\boldsymbol{x}^*$. Suppose that the solver prints that $\boldsymbol{x}^*$ is a local minimum, and you can confirm (based on your experience with the problem) that this is true. But you can also confirm that the vector $\boldsymbol{\lambda}^* \in \mathbb{R}^\ell$ which is also printed does not satisfy the KKT conditions together with $\boldsymbol{x}^*$. Explain how this can be the case.

(3p) Question 5

(modelling)

Sudoku is an old Japanese number puzzle that has grown in popularity recently. Some claim that it deals with numbers but has nothing to do with mathematics; however, this is obviously not true. The rules for a Sudoku puzzle are simple. We start with a 9×9-matrix $\boldsymbol{A}$ which is partitioned into nine 3×3-matrices $\boldsymbol{A}_1, \dots, \boldsymbol{A}_9$. Certain entries of $\boldsymbol{A}$ contain numbers from the set $\{1, \dots, 9\}$. An example of such a pre-assignment is shown below.

<u>The Sudoku problem is now to:</u>
Fill the empty entries of the matrix $\boldsymbol{A}$ such that
– each row of $\boldsymbol{A}$
– each column of $\boldsymbol{A}$
– each 3×3-submatrix $\boldsymbol{A}_1, \dots, \boldsymbol{A}_9$
contains every number of $\{1, \dots, 9\}$ exactly once.

Your task is to model an *integer linear program*, that is, a program that if all integral requirements would be relaxed is a linear program (the objective function and the constraints are all described by linear functions). This model shall, given

		5						3
				4	6			
		7						2
	1				3		6	9
	4		6		9		5	
9	8		2				7	
2						9		
			8	1				
6						4		

Figure 1: Example of a Sudoku problem.

a set of pre-assigned numbers, give a feasible solution to the Sudoku problem, and if this is not possible it should give a solution that minimizes the number of pre-assignments that cannot be fulfilled for the Sudoku problem to be solved.

Hint: Assume that the pre-assigned entries are collected in a set N whose size is n and is of the form $N = \{ (i, j, k) \mid \text{the number } k \text{ is pre-assigned to row } i, \text{ column } j \}$.

Question 6

(definitions)

(1p) a) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a given function. Define the property that f is a *strictly convex* function on $\mathbb{R}^n$.

(1p) b) Define the *termination criterion* that determines an *unbounded* LP solution.

(1p) c) Define the concept of an *extreme point* of a convex set.

Question 7

(Lagrangian duality for equality constrained problems)

Consider the primal problem

$$f^* := \underset{x}{\text{infimum}}\ f(\boldsymbol{x}), \tag{1a}$$
$$\text{subject to } \boldsymbol{x} \in X, \tag{1b}$$
$$h_j(\boldsymbol{x}) = 0, \qquad j = 1, \dots, \ell, \tag{1c}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h_j : \mathbb{R}^n \to \mathbb{R}$ $(j = 1, \dots, \ell)$ are continuous functions, and $X \subseteq \mathbb{R}^n$ is closed.

For an arbitrary vector $\boldsymbol{\lambda} \in \mathbb{R}^\ell$, we define the *Lagrange function*

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) := f(\boldsymbol{x}) + \sum_{j=1}^{\ell} \lambda_j h_j(\boldsymbol{x}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{h}(\boldsymbol{x}). \tag{2}$$

Let

$$q(\boldsymbol{\lambda}) := \underset{x \in X}{\text{infimum}}\ L(\boldsymbol{x}, \boldsymbol{\lambda}) \tag{3}$$

be the *Lagrangian dual function*, defined by the infimum value of the Lagrange function over X when we have Lagrangian relaxed the explicit constraints with multiplier values $\boldsymbol{\lambda}$; the *Lagrangian dual problem* is to

$$\underset{\boldsymbol{\lambda}}{\text{maximize}}\ q(\boldsymbol{\lambda}), \tag{4a}$$
$$\text{subject to } \boldsymbol{\lambda} \in \mathbb{R}^\ell. \tag{4b}$$

We suppose that X is non-empty, closed and bounded so that the "infimum" can be replaced by "minimum" in (1), (3), and both the primal and dual objective functions in (1) and (4) attain their optimal values. Moreover, then q is finite, continuous, and concave on the whole set $\mathbb{R}^\ell$, so that (4) is a "well-behaved" convex problem.

(1p) a) Let $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^\ell$. Define a subgradient, $\bar{\boldsymbol{\gamma}}$, to q at $\bar{\boldsymbol{\lambda}}$. Also define the entire subdifferential, $\partial q(\bar{\boldsymbol{\lambda}})$ to q at $\bar{\boldsymbol{\lambda}}$.

(1p) b) Let $\boldsymbol{x}(\bar{\boldsymbol{\lambda}})$ be an optimal solution to the Lagrangian subproblem defining the value of $q(\bar{\boldsymbol{\lambda}})$. Show that the vector $\boldsymbol{h}(\boldsymbol{x}(\bar{\boldsymbol{\lambda}})) \in \partial q(\bar{\boldsymbol{\lambda}})$; that is, show that the vector $\boldsymbol{h}(\boldsymbol{x}(\bar{\boldsymbol{\lambda}}))$ of constraint function values at the subproblem solution $\boldsymbol{x}(\bar{\boldsymbol{\lambda}})$ is a subgradient of q at $\bar{\boldsymbol{\lambda}}$.

(1p) c) Show that if, for some $\boldsymbol{\lambda}^* \in \mathbb{R}^\ell$, it holds that $\mathbf{0}^\ell \in \partial q(\boldsymbol{\lambda}^*)$ then $\boldsymbol{\lambda}^*$ is globally optimal in the Lagrangian dual problem (4).

Good luck!

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 06–03–06
Examiner: Michael Patriksson

Question 1

(the Simplex method)

(2p) a) After adding two slack variables, a BFS cannot be found directly. We create the phase I problem through an added artificial variable a_1 in the second linear constraint; the value of a_1 is to be minimized. We use the BFS based on the variable pair (s_1, a_1) as the starting BFS for the phase I problem. One iteration with the simplex methods gives the optimal basis $x_B = (s_1, x_3)^T$, which is a BFS for the original problem.

Starting phase II with this BFS, in the first iteration x_1 is the only variable with negative reduced cost and is picked as the incoming variable. The minimum ratio test shows that s_1 should leave the basis. In the next iteration, all reduced costs are > 0 and we conclude that we have found a unique optimal BFS $(x_1, x_3)^T = \frac{1}{5}(24, 13)^T$. The corresponding optimal objective value is $z^* = -\frac{11}{5}$.

(1p) b) Strong duality guarantees that if one of the primal or dual problem has an optimal solution then so does the other. Hence, the answer is yes.

(3p) Question 2

(The Karush–Kuhn–Tucker conditions)

See the proof of Theorem 5.25.

Question 3

(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.

(1p) a) False.

(1p) b) False.

(1p) c) Yes. The facts imply that the feasible set "expands" in a direction which we are both interested (dual variable positive) and able (non-degeneracy) to follow.

Question 4

(nonlinear programming)

(1p) a) The problem is convex when f is convex and each function h_j is affine.

(2p) b) One possibility is that no CQ is satisfied at $\boldsymbol{x}^*$.

It could also be the case that there exists a local minimum which is also a KKT point close to $\boldsymbol{x}^*$. (For numerical reasons the algorithm has terminated prematurely, and so $\boldsymbol{x}^*$ may only be near-optimal, which means that the KKT conditions cannot be satisfied exactly.)

Question 5

(modelling)

Introduce the variables

$$x_{ijk} = \begin{cases} 1 \text{ if number k is chosen for the entry on row i, column j} \\ 0 \text{ else} \end{cases}, \tag{1}$$

which is defined for $i, j, k = 1, \dots, 9$. We need the constraints

$$\sum_i x_{ijk} = 1, \quad \forall j, k \tag{2}$$

$$\sum_j x_{ijk} = 1, \quad \forall i, k \tag{3}$$

$$\sum_k x_{ijk} = 1, \quad \forall i, j \tag{4}$$

$$\sum_{i=3p-2}^{3p} \sum_{j=3p-2}^{3p} x_{ijk} = 1, \quad \forall k,\ p = 1, 2, 3 \tag{5}$$

$$x_{ijk} \in \{0, 1\}, \quad \forall i, j, k \tag{6}$$

Equation 2 makes sure that each column contains each number, equation 3 that each row contains each number, equation 5 that each submatrix contains each number and equation 4 that each entry has exactly one number assigned to it. Equation 6 is a logic equation, saying that either is a number picked or not.

The objective function is

$$\min z = n - \sum_{ijk \in N} x_{ijk}, \tag{7}$$

where $z^* = 0$ means that the Sudoku problem could be solved with all the pre-assignments kept.

Question 6

(definitions)

(1p) a) See Definition 3.33.

(1p) b) See Step 2 on page 229.

(1p) c) See Definition 3.11.

Question 7

(Lagrangian duality for equality constrained problems)

(1p) a) At $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^\ell$, a subgradient $\bar{\boldsymbol{\gamma}}$ of the concave function q is such that

$$q(\boldsymbol{\lambda}) \leq q(\bar{\boldsymbol{\lambda}}) + \bar{\boldsymbol{\gamma}}^{\mathrm{T}}(\boldsymbol{\lambda} - \bar{\boldsymbol{\lambda}}), \qquad \boldsymbol{\lambda} \in \mathbb{R}^\ell.$$

The subdifferential $\partial q(\bar{\boldsymbol{\lambda}})$ to q at $\bar{\boldsymbol{\lambda}}$ is the convex hull of all the subgradients $\bar{\boldsymbol{\gamma}}$ of q at $\bar{\boldsymbol{\lambda}}$.

(1p) b) Let $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^\ell$, and $\bar{\boldsymbol{x}} \in X(\bar{\boldsymbol{\lambda}})$. Then,

$$\begin{aligned} q(\boldsymbol{\lambda}) = & \operatorname*{infimum}_{y \in X} L(\boldsymbol{y}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x}) \\ &= f(\boldsymbol{x}) + \bar{\boldsymbol{\lambda}}^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x}) + (\boldsymbol{\lambda} - \bar{\boldsymbol{\lambda}})^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x}) \leq q(\bar{\boldsymbol{\lambda}}) + (\boldsymbol{\lambda} - \bar{\boldsymbol{\lambda}})^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x}), \end{aligned}$$

which implies that $\boldsymbol{h}(\boldsymbol{x}) \in \partial q(\boldsymbol{\lambda})$.

(1p) c) Suppose that $\mathbf{0}^\ell \in \partial q(\boldsymbol{\lambda})$. From b) then follows that $q(\boldsymbol{\lambda}) \geq q(\bar{\boldsymbol{\lambda}})$ for all $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^\ell$.
