Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 08–03–25
Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) To transform the problem to standard form, first replace the free variable x_1 with the non-negative variables x_1^+ and x_1^- such that $x_1 = x_1^+ - x_1^-$. Then change sign on the inequality constraint and substract a slack variable $s_1 \geq 0$. A BFS cannot be found directly, hence begin with a phase 1 problem using artificial variables $a_1, a_2 \geq 0$ in both constraints. The objective is to minimize $a_1 + a_2$. Start with the BFS given by (a_1, a_2) in the basis. In the first iteration of the simplex algorithm, x_1^- is the only variable with a negative reduced cost (-4), and is therefore the only eligable incoming variable. The minimum ratio test shows that a_2 should leave the basis. In the next iteration, s_1 is the only variable with a negative reduced cost $(-\frac{1}{3})$ and is chosen as the incoming variable. The minimum ratio test shows that x_3 should leave. No artificial variables are left in the basis, and we can proceed to phase 2.

The reduced costs in the first iteration of the phase 2 problem are

$$\tilde{c}^{\mathrm{T}}_{(x_1^+, x_2, x_3)} = (0, 1, 1) \geq \mathbf{0},$$

and thus the optimality condition is fulfilled for the current basis. We have $\boldsymbol{x}_B^* = (5, 2)^{\mathrm{T}}$, or, in the original variables, $\boldsymbol{x}^* = (x_1, x_2, x_3)^* = (-2, 0, 0)^{\mathrm{T}}$, with the optimal value $z^* = 2$.

(1p) b) The dual to the LP is given by

$$\begin{array}{llrcl}
\text{maximize} & w = & 2\,x_1 - \;\; y_2, & & \\
\text{subject to} & & -\;\; y_1 + 3y_2 & = & -1, \\
 & & -\,2y_1 + \;\; y_2 & \leq & -1, \\
 & & -\;\; y_1 \qquad & \leq & 0, \\
 & & y_1 \qquad & \in & \mathbb{R} \text{ (free)}, \\
 & & y_2 & \leq & 0.
\end{array}$$

The primal problem has an optimal solution. Then, from strong duality, so does the dual problem. Add a slack variable $y_3 \geq 0$ in the second constraint and let the dual optimal solution be $\boldsymbol{y}^*$. There cannot exist a solution $\boldsymbol{u}$ to the given system, since if that would be the case, then $\boldsymbol{y}^* + \boldsymbol{u}$ would be feasible in the dual with a larger objective value (from the third row in the system). This is a contradiction to $\boldsymbol{y}^*$ being optimal.

Question 2

(modelling)

(1p) a) One possibility is the following formulation:

$$
\begin{aligned}
\min \quad & f(\boldsymbol{x}, \boldsymbol{y}), & & \\
\text{s.t.} \quad & x_j(1 - x_j) = 0, & & j = 1, \dots, n, \\
& \boldsymbol{Ax} \leq \boldsymbol{b}, & & \\
& y_i \geq 0, & & i = 1, \dots, m.
\end{aligned} \tag{NLP}
$$

(2p) b) Two problems must be solved. The optimal solution to MIP is given by the solution to the problem with the least optimal value. That is, the optimal value of MIP is

$$z^* = \min\{z_0^*, z_1^*\},$$

where

$$
\begin{aligned}
z_0^* = \min \quad & f(x, \boldsymbol{y}), & & \\
\text{s.t.} \quad & x = 0, & & \\
& \boldsymbol{Ax} \leq \boldsymbol{b}, & & \\
& y_i \geq 0, & & i = 1, \dots, m,
\end{aligned} \tag{P^0}
$$

and

$$
\begin{aligned}
z_1^* = \min \quad & f(x, \boldsymbol{y}), & & \\
\text{s.t.} \quad & x = 1, & & \\
& \boldsymbol{Ax} \leq \boldsymbol{b}, & & \\
& y_i \geq 0, & & i = 1, \dots, m.
\end{aligned} \tag{P^1}
$$

Question 3

(topics in Lagrangian duality)

(1p) a) See The Book, Theorem 6.4.

(1p) b) Under the assumptions on X, for any vector $\boldsymbol{\mu} \in \mathbb{R}^m$ the function $L(\cdot, \boldsymbol{\mu})$ is weakly coercive with respect to X (see The Book, Definition 4.5). By the continuity assumptions on f and g_i, $i = 1, \dots, m$, $L(\cdot, \boldsymbol{\mu})$ is also continuous. Hence, Weierstrass' Theorem 4.7 applies.

(1p) c) $x^* = 0$; the dual problem has no optimal solution; however, $f^* = q^* = 0$, whence the duality gap $\Gamma = 0$.

(3p) Question 4

(complementarity slackness theorem)

We first establish that if the system (3) is satisfied at $(\boldsymbol{x}, \boldsymbol{y})$ then the pair $(\boldsymbol{x}, \boldsymbol{y})$ is primal–dual optimal in (1), (2). By assumption, $\boldsymbol{x}$ (respectively, $\boldsymbol{y}$) is a feasible solution to the primal (respectively, dual) problem. By the Weak Duality Theorem 10.5, then, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$. The system (3) implies that in fact equality holds. This immediately, by the Corollary 10.6 to the Weak Duality Theorem, implies that the pair $(\boldsymbol{x}, \boldsymbol{y})$ must be optimal.

Suppose then that the pair $(\boldsymbol{x}, \boldsymbol{y})$ is primal–dual optimal in (1), (2). Then, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ holds, by the Strong Duality Theorem. In the string of inequalities

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq \boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} \leq \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$$

provided by the Weak Duality Theorem 10.5, equality then must hold throughout. From the resulting two equalities then follow (3).

Question 5

(quadratic programming)

(1p) a) The KKT-conditions are:

$$2\boldsymbol{H}\boldsymbol{x} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{0}^n$$

The problem is a convex problem with linear constraints, so a feasible solution which fulfills the KKT-conditions is a global optimal solution. Since $\boldsymbol{H}$ is postive definite and hence invertible, we have:

$$\boldsymbol{x}^* = -\tfrac{1}{2}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda},$$

and

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b} \Rightarrow -\tfrac{1}{2}\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{b}.$$

Since $\boldsymbol{A}$ has full row rank, $\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}}$ is invertible, and so

$$\boldsymbol{\lambda}^* = -2(\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{b}.$$

(1p) b) The dual function is found by minimizing the Lagrangian for each $\boldsymbol{\lambda}$. So

$$\boldsymbol{x}^*(\boldsymbol{\lambda}) = -\tfrac{1}{2}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda},$$

which gives

$$q(\boldsymbol{\lambda}) = L(\boldsymbol{x}^*(\boldsymbol{\lambda}), \boldsymbol{\lambda}) = -\tfrac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b}.$$

In the dual problem, we want to maximize the dual function. Since we have equality constraints in the primal, we have no bounds on the dual variables:

$$\text{maximize} \quad q(\boldsymbol{\lambda}) := -\tfrac{1}{4}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b}. \qquad \text{(Dual QP)}$$

The dual problem is a convex unconstrained problem, and a dual optimal solution is therefore found by setting the gradient of q to zero, which yields $\boldsymbol{\lambda}^* = -2(\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{b}$.

(1p) c) Let $\hat{\boldsymbol{x}}$ be a feasible solution, i.e., $\boldsymbol{A}\hat{\boldsymbol{x}} = \boldsymbol{b}$. Then $\boldsymbol{x} = \hat{\boldsymbol{x}} + \boldsymbol{Z}\boldsymbol{p}$, $\boldsymbol{p} \in \mathbb{R}^{n-m}$ is also a feasible solution, and (QP) is equivalent to:

$$\underset{\boldsymbol{p}\in\mathbb{R}^{n-m}}{\text{minimize}} \; \boldsymbol{p}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{Z}\boldsymbol{p} + 2\boldsymbol{p}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{H}\hat{\boldsymbol{x}} + \text{const.}$$

This is an unconstrained problem with a pos. semidef. Hessian, and hence it is a convex problem. A local optimal solution is a global optimal solution.

(3p) Question 6

(The Frank-Wolfe algorithm)

Since the objective function is nonconvex, we cannot provide any lower bounds from the subproblem solutions. An upper bound is $f(\boldsymbol{x}^0) = 0.5$. At $\boldsymbol{x}^0 = (1,1)^{\mathrm{T}}$, $\nabla f(\boldsymbol{x}^0) = (1,0)^{\mathrm{T}}$; $\boldsymbol{y}^0 = (0,1)^{\mathrm{T}}$; $\text{argmin}_{\alpha\in[0,1]}\,\varphi(\alpha) = 1$, where $\varphi(\alpha) = f(\boldsymbol{x}^0 + \alpha(\boldsymbol{y}^0 - \boldsymbol{x}^0))$; $\boldsymbol{x}^1 = (0,1)^{\mathrm{T}}$; $\nabla f(\boldsymbol{x}^1) = (0,0)^{\mathrm{T}}$. A new upper bound is $f(\boldsymbol{x}^1) = 0$. The vector $\boldsymbol{x}^1$ is a KKT-point (set all Lagrange multipliers to zero). It is not a local minimum, however, since for example $\boldsymbol{x}(t) = \boldsymbol{x}^1 + (t, 4t)^{\mathrm{T}}$ is feasible for $0 \le t \le 0.25$, and for $t > 0$, $f(\boldsymbol{x}(t)) = -\frac{15}{2}t^2 < f(\boldsymbol{x}^1)$. [There are two global minima: $\boldsymbol{x}^* = (0.25, 0)^{\mathrm{T}}$ and $\boldsymbol{x}^* = (0.25, 2)^{\mathrm{T}}$.]

Question 7

(nonlinear optimization solves interesting problems)

(1p) a) Let $\varepsilon > 0$ be any small enough number. The optimization problem is to find

$$\begin{aligned} f^* = \underset{(x,y,z,n)\in\mathbb{R}^4_+}{\text{minimum}} \quad & f(x,y,z,n) := (x^n + y^n - z^n)^2, \\ \text{subject to} \quad & \sin \pi x = \sin \pi y = \sin \pi z = \sin \pi n = 0, \\ & xyz \geq \varepsilon, \\ & n \geq 3. \end{aligned}$$

If $f^* > 0$ then Fermat's Last Theorem has been proved. (Which it already has by other means.)

(1p) b) Consider the problem to

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \; & f(\boldsymbol{x}) := \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}, \\ \text{subject to} \; & \|\boldsymbol{x}\| = 1. \end{aligned}$$

An optimal solution, say $\boldsymbol{x}^*$, exists due to Weierstrass' Theorem, as the sphere is non-empty, closed and bounded. For each non-zero vector $\boldsymbol{x} \in \mathbb{R}^n$, the vector $\|\boldsymbol{x}\|^{-1}\boldsymbol{x}$ is a feasible solution; hence, $\|\boldsymbol{x}\|^{-2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geq (\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}^* =: c$.

(1p) c) Choose $\boldsymbol{y} \in \mathbb{R}^n$ aribtrarily. To prove existence and uniqueness of a solution to the equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$, consider the minimization of $f(\boldsymbol{x}) := \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}$ over $\boldsymbol{x} \in \mathbb{R}^n$.

The function f is coercive on $\mathbb{R}^n$, whence Weierstrass' Theorem applies; the problem has an optimal solution. As it is unconstrained, we know that stationarity is a necessary condition, so we set the gradient of f to zero: $\nabla f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0}^n$, and so we know that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ holds. To establish uniqueness, we may observe that $\boldsymbol{A}\boldsymbol{x}^1 = \boldsymbol{A}\boldsymbol{x}^2 = \boldsymbol{y}$ implies that $\boldsymbol{A}(\boldsymbol{x}^1 - \boldsymbol{x}^2) = \boldsymbol{0}^n$ and hence that $(\boldsymbol{x}^1 - \boldsymbol{x}^2)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}^1 - \boldsymbol{x}^2) = 0$. By positive definiteness this implies that $\boldsymbol{x}^1 = \boldsymbol{x}^2$. We are done.