Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 13–12–17
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:

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

Phase I
We introduce an artificial variable a and formulate our Phase I problem.

$$\begin{aligned}
\text{minimize} \quad & z = a \\
\text{subject to} \quad & x_1 + 3x_2 - x_3 + s_1 = 5, \\
& 2x_1 - x_2 + 2x_3 - s_2 + a = 2, \\
& x_1, \; x_2, \; x_3, \; s_1, \; s_2, \; a \geq 0.
\end{aligned}$$

We now have a starting basis (s_1, a). Calculating the reduced costs we obtain $\tilde{\mathbf{c}}_N = (-2, 1, -2, 1)^{\mathrm{T}}$, meaning that x_1 or x_3 should enter the basis. We choose x_3. From the minimum ratio test, we get that a should leave the basis. This concludes Phase I and we now have the basis (s_1, x_3).

Phase II
Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (2, -\frac{1}{2}, \frac{1}{2})^{\mathrm{T}}$. meaning that x_2 should enter 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 $\tilde{\mathbf{c}}_N = (\frac{12}{5}, \frac{1}{5}, \frac{2}{5})^{\mathrm{T}} \geq 0$, meaning that the current basis is optimal. The optimal solution is thus

$$(x_1, x_2, x_3, s_1, s_2)^{\mathrm{T}} = (0, \frac{12}{5}, \frac{11}{5}, 0, 0, 0)^{\mathrm{T}},$$

which in the original variables means $(x_1, x_2, x_3)^{\mathrm{T}} = (0, \frac{12}{5}, \frac{11}{5})^{\mathrm{T}}$ with optimal objective value $f^\star = -\frac{1}{5}$.

(1p) b) Calculating the reduced costs of the modified problem for the optimal basis of the original problem, we obtain $\tilde{\mathbf{c}}_N = (\frac{12}{5}, \frac{1}{5}, \frac{2}{5}, \frac{7}{10})^{\mathrm{T}} \geq 0$ meaning that the the optimal basis from the original problem gives the optimal solution of the modified problem $(x_1, x_2, x_3, x_4)^{\mathrm{T}} = (0, \frac{12}{5}, \frac{11}{5}, 0)^{\mathrm{T}}$ with optimal objective value $f^\star = -\frac{1}{5}$.

Question 2

(nonlinear programming)

(1p) a) As $\nabla f(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{c}$, we have that $\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p} = \boldsymbol{p}^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{c})$. With $\boldsymbol{x} = (-3, 4)^{\mathrm{T}}$ we hence have that descent is obtained whenever $\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p} < 0$, i.e. whenever $p_1(-3-c_1)+p_2(4-c_2) < 0$. Further if $\boldsymbol{p} \neq \boldsymbol{0}$ and $f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p} \geq 0$, by strict convexity of f we have, for any $\delta > 0$, that $f(\boldsymbol{x} + \delta\boldsymbol{p}) > f(\boldsymbol{x}) + \delta\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p} \geq f(\boldsymbol{x})$, so $\boldsymbol{p}$ is a descent direction to f at $\boldsymbol{x} = (-3, 4)^{\mathrm{T}}$ precisely when $p_1(-3 - c_1) + p_2(4 - c_2) < 0$.

(2p) b) With the set-up considered we will, for a given penalty parameter value k (a non-negative integer) consider the following penalty function to be minimized over $\mathbb{R}$:

$$P_k(x) := \begin{cases} x^2 + k(1 - x), & x < 1, \\ x^2, & x \geq 1, \end{cases}$$

The minimizer x_k^* of P_k is at $x = \frac{1}{2}$ for $k = 1$ and at $x = 1$ for all positive integers $k \geq 2$. The latter is also the optimal solution to the problem.

(3p)

Question 3

(characterization of convexity in C^1)

This is Theorem 3.61(a).

Question 4

(modelling)

(2p) a) Let x_t denote the number of units purchased from the producer on day t, and let y_t denote the number of units in the storage at the beginning of time t. Then the model is

$$
\begin{aligned}
\text{minimize} \quad & \sum_{t=1}^{7} (c_t x_t + g y_t), \\
\text{subject to} \quad & x_t + y_t \geq d_t, && t = 1, \ldots, 7, \\
& y_{t+1} = y_t + x_t - d_t, && t = 1, \ldots, 6, \\
& y_1 = 0, \\
& y_t \leq M, && t = 1, \ldots, 7, \\
& x_t, y_t \geq 0, && t = 1, \ldots, 7.
\end{aligned}
$$

(1p) b) We now introduce variables x_t^{high} denoting the number of units purchased on day t for the higher price c_t^{high}, and x_t^{low} denoting the number of units purchased on day t for the lower price c_t. Now the model is

$$
\begin{aligned}
\text{minimize} \quad & \sum_{t=1}^{7} \left(c_t^{\text{high}} x_t^{\text{high}} + c_t x_t^{\text{low}} + g y_t \right), \\
\text{subject to} \quad & x_t + y_t \geq d_t, && t = 1, \ldots, 7, \\
& y_{t+1} = y_t + x_t - d_t, && t = 1, \ldots, 6, \\
& y_1 = 0, \\
& y_t \leq M, && t = 1, \ldots, 7, \\
& x_t = x_t^{\text{high}} + x_t^{\text{low}}, && t = 1, \ldots, 7, \\
& x_t^{\text{low}} \leq K, && t = 1, \ldots, 7, \\
& x_t, y_t \geq 0, && t = 1, \ldots, 7.
\end{aligned}
$$

Question 5

(true or false)

(1p) a) False. Counter example: $f = |x|$, $x = 0$, and $p = 1$. Then p is a subgradient to f at $x = 0$, but it is not a descent direction.

(1p) b) True. The claim follows directly from weak duality.

(1p) c) True. For all $\boldsymbol{x} \in S$, we can choose $\boldsymbol{y} = \boldsymbol{x}$ in the minimization, implying that the value of the infimum must be smaller than or equal to zero.

Question 6

(KKT conditions)

(2p) a)

Notice that the constraint $(x_1 + x_2 - 4)^2 \geq 1$ is active precisely when $x_1 + x_2 - 4 = \pm 1$. The feasible set can thus be drawn as two disjoint triangles with extreme points in $(0,0),(0,3),(3,0)$ and $(4,4),(4,1),(1,4)$, respectively.

The level curves of the objective function are circles centred in $(2,2)$, so the negative gradient of the objective function at $\mathbf{x}$ lies along the line from $(2,2)$ to $\mathbf{x}$. This allows us to draw the figure 1. Searching for points where the negative objective function lies in the normal cone, i.e., $-\nabla f(\mathbf{x}) \in N_S(\mathbf{x})$, we graphically find the KKT-points indicated in the figure. Thus the KKT points are $\{(0,0),(2,0),(3,0),(4,1),(4,2),(4,4),(2,4),(1,4),(0,3),(0,2)\}$.

(1p) b) To motivate logically we need to establish two claims.
Claim 1: The problem has *some* optimal solution.
Claim 2: Any (locally) optimal solution is a KKT-point.

To establish Claim 1, we note that the objective function and all constraint functions are continuous, so the feasible set S is closed. Further S is clearly bounded, due to the constraints $0 \leq x_i \leq 4$. Hence Weierstrass' Theorem establishes the claim.

To establish Claim 2, we recall that any locally optimal is a KKT-point if some constraint qualification holds. Looking at figure 1, we can note that the gradients of the active constraints are linearly independent in each point, hence LICQ holds.

Question 7

(linear programming duality and optimality)

(1p) a) For problem a), let y_{ij} be the dual variable associated with constraint $t_j -$

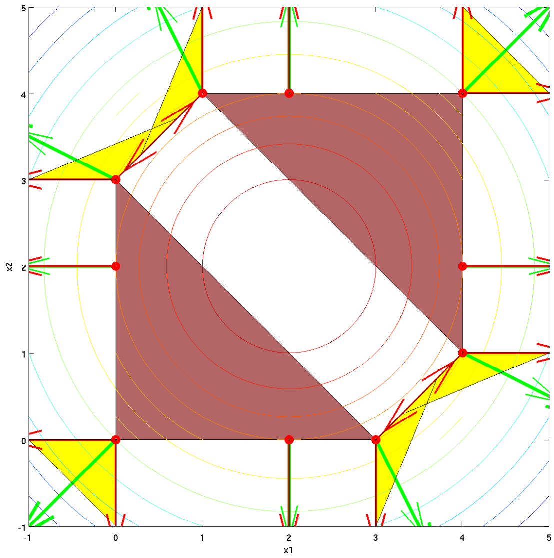

Figure 1: Feasible set and level curves of the objective function. Green arrows indicate the negative objective function gradient, red arrows indicate gradients of active constraints. The normal cones are indicated in yellow.

$t_i \geq T_i$ corresponding to edge (i, j), from node i to node j. Then the dual problem can be written as

$$
\begin{array}{lrrrrrrrrrrl}
\underset{y_{12},\, y_{13},\, y_{23},\, y_{24},\, y_{34}}{\text{maximize}} & & T_1 y_{12} & + & T_1 y_{13} & + & T_2 y_{23} & + & T_2 y_{24} & + & T_3 y_{34} & \\
\text{subject to} & - & y_{12} & - & y_{13} & & & & & & & = -1 \\
 & & y_{12} & & & - & y_{23} & - & y_{24} & & & = 0 \\
 & & & & y_{13} & + & y_{23} & & & - & y_{34} & = 0 \\
 & & & & & & & + & y_{24} & + & y_{34} & = 1 \\
 & & y_{12}, & & y_{13}, & & y_{23}, & & y_{24}, & & y_{34} & \geq 0.
\end{array}
$$

(1p) b) For problem b), let t_i^* be the optimal starting times in the primal problem. The reason behind the given expressions for the primal optimal solutions is as follows:

$$
\begin{aligned}
t_2^* &= t_1^* + T_1 = t_1^* + 1 \\
t_3^* &= \max\{t_1^* + T_1, t_2^* + T_2\} = \max\{t_1^* + 1, t_1^* + 1 + 2\} = t_1^* + 3 \\
t_4^* &= \max\{t_2^* + T_2, t_3^* + T_3\} = \max\{t_1^* + 1 + 2, t_1^* + 3 + 1\} = t_1^* + 4
\end{aligned}
$$

Therefore, the optimal objective value of the primal problem is $t_4^* - t_1^* = 4$. Notice that the precedence constraints associated with edges $(1, 2)$, $(2, 3)$, $(3, 4)$ are active but those with edges $(1, 3)$ and $(2, 4)$ are not active.

For the dual problem, consider the $0-1$ binary valued dual optimal solution candidate according to the rule that $y^*_{ij} = 1$ if and only if edge (i, j) is on the path from node 1 to node 4 with the maximum sum of edge weights. This path is $(1, 2) \to (2, 3) \to (3, 4)$. Thus, $y^*_{12} = y^*_{23} = y^*_{34} = 1$ and $y^*_{13} = y^*_{24} = 0$. These dual variables are feasible, and the corresponding dual objective value is 4 which is the same as the optimal primal objective value. Therefore, the weak duality theorem implies that y^*_{ij} are indeed dual optimal.

For your information, the dual problem has the interpretation of a maximum cost flow problem where one unit of "flow" is shipped from the source (node 1) to the sink (node 4). The total supply to node 1 is one unit (i.e., the first constraint in the dual problem), and the total demand at node 4 is one unit (i.e., the last constraint in the dual problem). In addition, for node 2 and node 3, the total incoming flow is equal to the total outgoing flow (i.e., flow is conserved). The dual problem seeks to route the one unit of flow through the network in order to maximize the cost in the dual objective function. Because of the integer-valued supply and demand, the maximum cost flow problem amounts to finding the path of the maximum sum of edge weights from the source to the sink.

(1p) c) For problem c), the primal and dual optimal solutions can be verified to satisfy

$$y^*_{ij}(t^*_j - t^*_i - T_i) = 0, \quad \text{for all edges } (i, j).$$

These are the complementary slackness optimality conditions. In particular, for edges $(1, 2)$, $(2, 3)$ and $(3, 4)$ where $y^*_{ij} = 1$, the corresponding primal precedence constraints are active (i.e., $t^*_j - t^*_i - T_i = 0$). On the other hand, for edges $(1, 3)$ and $(2, 4)$ where the primal precedence constraints are not active, the corresponding dual variables must be zero.