Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 07–12–17
Examiner: Michael Patriksson

Question 1

(The Simplex method)

(2p) a) By introducing a slack variable x_5 and two artificial variables a_1 and a_2, we get the Phase I problem to

$$\begin{array}{rrrrrrrrcl} \text{minimize} \quad w = & & & & & & a_1 & +a_2 & & \\ \text{subject to} & x_1 & & -x_3 & & & +a_1 & & = & 3\,, \\ & x_1 & -x_2 & & -2x_4 & & & +a_2 & = & 1\,, \\ & 2x_1 & & & +x_4 & +x_5 & & & = & 7\,, \\ & x_1\,, & x_2\,, & x_3\,, & x_4\,, & x_5\,, & a_1\,, & a_2 & \geq & 0\,. \end{array}$$

Let $\boldsymbol{x}_B^{\mathrm{T}} = (a_1, a_2, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (x_1, x_2, x_3, x_4)$ be the initial basic and nonbasic vector. The reduced costs of the nonbasic variables are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (-2, 1, 1, 2),$$

which means that x_1 is the entering variable. Further, we have

$$\boldsymbol{B}^{-1}\boldsymbol{b} = (3, 1, 7)^{\mathrm{T}},$$

$$\boldsymbol{B}^{-1}\boldsymbol{N}_1 = (1, 1, 2)^{\mathrm{T}},$$

which gives

$$\operatorname{argmin}_{j:(B^{-1}N_1)_j>0} \frac{(\boldsymbol{B}^{-1}\boldsymbol{b})_j}{(\boldsymbol{B}^{-1}\boldsymbol{N}_1)_j} = 2,$$

so a_2 is the leaving variable. The new basic and nonbasic vectors are $\boldsymbol{x}_B^{\mathrm{T}} = (a_1, x_1, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (a_2, x_2, x_3, x_4)$, and the reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (2, -1, 1, -2),$$

so x_4 is the entering variable, and

$$\boldsymbol{B}^{-1}\boldsymbol{b} = (2, 1, 5)^{\mathrm{T}},$$

$$\boldsymbol{B}^{-1}\boldsymbol{N}_4 = (2, -2, 5)^{\mathrm{T}},$$

which gives

$$\operatorname{argmin}_{j:(B^{-1}N_4)_j>0} \frac{(\boldsymbol{B}^{-1}\boldsymbol{b})_j}{(\boldsymbol{B}^{-1}\boldsymbol{N}_4)_j} = 1,$$

and thus a_1 is the leaving variable. The new basic and nonbasic vectors are $\boldsymbol{x}_B^{\mathrm{T}} = (x_4, x_1, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (a_2, x_2, x_3, a_1)$, and the reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (1, 0, 0, 1),$$

so $\boldsymbol{x}_B^{\mathrm{T}} = (x_4, x_1, x_5)$ is an optimal basic feasible solution of the Phase I problem. Since $w^* = 0$, $\boldsymbol{x}_B$ is a basic feasible solution of the Phase II problem to

$$\begin{array}{lrrrrrcl} \text{minimize} & z = 2x_1 & & & & & & \\ \text{subject to} & x_1 & & -x_3 & & & = & 3, \\ & x_1 & -x_2 & & -2x_4 & & = & 1, \\ & 2x_1 & & & +x_4 & +x_5 & = & 7, \\ & x_1, & x_2, & x_3, & x_4, & x_5 & \geq & 0. \end{array}$$

If $\boldsymbol{x}_B^{\mathrm{T}} = (x_4, x_1, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (x_2, x_3)$, we get the reduced costs

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (0, 2).$$

This means that $\boldsymbol{x}_B$ is an optimal basic feasible solution for the Phase II problem, and we are done! $\boldsymbol{x}^* = (3, 0, 0, 1)^{\mathrm{T}}$ and $z^* = 6$.

(1p) b) If the primal is infeasible, the dual cannot have an optimal solution. Thus it is either infeasible or unbounded.

Question 2

(the KKT conditions)

(1p) a) See the Book, system (5.9).

(1p) b) The vector $\boldsymbol{x}^1$ satisfies the KKT conditions (5.9).

(1p) c) Nothing. (Under the conditions given, there may be optimal solutions that do not satisfy the KKT conditions.)

Question 3

(short questions on different topics)

(1p) a) Yes it is. $(1, 0, 1, 0, 0)^{\mathrm{T}}$ is feasible and the columns of A corresponding to the positive entries are linearly independent.

(1p) b) By multiplying with $\boldsymbol{p}_k$ from the left we get

$$\boldsymbol{p}_k^{\mathrm{T}}(\nabla^2 f(\boldsymbol{x}_k) + \gamma_k \boldsymbol{I}^n)\boldsymbol{p}_k = -\boldsymbol{p}_k^{\mathrm{T}}\nabla f(\boldsymbol{x}_k).$$

Since γ_k is chosen such that $\nabla^2 f(\boldsymbol{x}_k) + \gamma_k \boldsymbol{I}^n$ is positive definite [that is, $\boldsymbol{u}^{\mathrm{T}}(\nabla^2 f(\boldsymbol{x}_k) + \gamma_k \boldsymbol{I}^n)\boldsymbol{u} > 0$ holds for all $\boldsymbol{u} \in \mathbb{R}^n \setminus \{\boldsymbol{0}^n\}$], it follows that $\boldsymbol{p}_k^{\mathrm{T}}\nabla f(\boldsymbol{x}_k) < 0$ and $\boldsymbol{p}_k$ is therefore a direction of descent.

(1p) c) It is not true. Consider for example the problem to

$$\begin{aligned} &\text{minimize} && x_1, \\ &\text{subject to} && x_1^2 + x_2^2 - 1 = 0, \\ & && x \in X = \{\, \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 + x_2 \geq 0 \,\}, \end{aligned}$$

which has the two local minima $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ and $\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, of which only the latter is a global minimum.

(3p) Question 4

(the separation theorem) See the Book, Theorem 4.28.

Question 5

(LP duality and derivatives)

(1p) a) If $v(\boldsymbol{b})$ is finite, then by LP duality, we have that

$$\begin{aligned} v(\boldsymbol{b}) := \underset{y \in \mathbb{R}^m}{\text{maximum}} \quad & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}, \\ \text{subject to} \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c}, \\ & \boldsymbol{y} \text{ free}. \end{aligned} \tag{1}$$

At least one maximum in (1) is attained at an extreme point of the dual polyhedron. Therefore, we can write $v(\boldsymbol{b}) = \text{maximum}_{k \in \mathcal{K}}\, \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_k$, where $\{\boldsymbol{y}_k\}_{k \in \mathcal{K}}$ is the (finite) set of extreme points of the dual polyhedron. The convexity of v follows simply by using the definition: for $\lambda \in (0,1)$ and arbitrary vectors $\boldsymbol{b}^1$ and $\boldsymbol{b}^2$ in $\mathbb{R}^m$ it holds that

$$\max_{k \in \mathcal{K}}\, [\lambda \boldsymbol{b}^1 + (1-\lambda)\boldsymbol{b}^2]^{\mathrm{T}}\boldsymbol{y}_k \leq \lambda \max_{k \in \mathcal{K}}\, (\boldsymbol{b}^1)^{\mathrm{T}}\boldsymbol{y}_k + (1-\lambda) \max_{k \in \mathcal{K}}\, (\boldsymbol{b}^2)^{\mathrm{T}}\boldsymbol{y}_k,$$

the inequality being a consequence of the added freedom of choice when separating the optimization problem on the left-hand side of the inequality with the two optimization problems in the right-hand side. Hence,

$$v(\lambda \boldsymbol{b}^1 + (1-\lambda)\boldsymbol{b}^2) \leq \lambda v(\boldsymbol{b}^1) + (1-\lambda)v(\boldsymbol{b}^2),$$

and we are done.

(2p) b) Consider the following inequality:

$$v(\boldsymbol{p}) \geq v(\boldsymbol{b}) + \boldsymbol{\xi}^{\mathrm{T}}(\boldsymbol{p} - \boldsymbol{b}), \qquad \forall \boldsymbol{p} \in \mathbb{R}^m,$$

where $\boldsymbol{\xi} \in \mathbb{R}^m$. This inequality is the definition of the vector $\boldsymbol{\xi}$ being a subgradient of the convex function v at $\boldsymbol{b}$; it in fact characterizes v as being convex, whenever it is sub-differentiable. Our task is to establish that this inequality holds when we let $\boldsymbol{\xi} = \boldsymbol{y}^*$. Since $v(\boldsymbol{b}) = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^*$ by assumption, the inequality reduces to stating that

$$v(\boldsymbol{p}) \geq \boldsymbol{p}^{\mathrm{T}}\boldsymbol{y}^*, \qquad \forall \boldsymbol{p} \in \mathbb{R}^m.$$

But this is true: by definition, $v(\boldsymbol{p})$ equals the supremum of $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{y}$ over all feasible vectors $\boldsymbol{y}$, and $\boldsymbol{y}^*$ is just one out of all the possible choices of dual feasible vectors.

Finally, differentiability of v at $\boldsymbol{b}$ is equivalent, given its convexity, to the existence of a unique subgradient of v at $\boldsymbol{b}$. From the above it is clear that if there is only one optimal solution to the problem (1) then that must also be the gradient of v at $\boldsymbol{b}$.

(3p) Question 6

(modelling) Introduce the variables:

x_i is 0 if element i is assigned to computer 1
and it is 1 if assigned to computer 2. $i = 1, \ldots, n$

y_k is 1 if edge k is between to elements assigned to different computers.
It is 0 otherwise. $k = 1, \ldots m$

The computing time for the elements is equal to

$$\max\left\{\frac{\eta}{\nu}\sum_{i=1}^{n} x_i,\ \frac{\eta}{\nu}\left(n - \sum_{i=1}^{n} x_i\right)\right\},$$

which can be modelled using an auxilary variable t and linear inequalities. The optimization problem reads:

$$
\begin{aligned}
\text{minimize} \quad z = \frac{\eta t}{\nu} + \frac{\rho}{\nu} \sum_{k=1}^{m} & \quad y_k \\
\text{subject to} \quad \sum_{i=1}^{n} x_i \leq & \quad t \\
n - \sum_{i=1}^{n} x_i \leq & \quad t \\
x_{E_{k,1}} - x_{E_{k,2}} \leq & \quad y_k \ , \ k = 1, \dots, m \\
x_{E_{k,2}} - x_{E_{k,1}} \leq & \quad y_k \ , \ k = 1, \dots, m \\
\boldsymbol{x} \in & \quad \mathbb{B}^n \\
\boldsymbol{y} \in & \quad \mathbb{B}^m \\
t \in & \quad \mathbb{R}
\end{aligned}
$$

(3p) Question 7

(Lagrangian Duality) Lagrangian relax the contraint to get

$$L(\boldsymbol{x}, \lambda) = -\lambda b + \sum_{i=1}^{n} x_i^2 + \lambda(\sum_{i=1}^{n} x_i - b).$$

L is differentiable and we find the Lagrangian dual function

$$q(\lambda) = \min_{x \in \mathbb{R}^n} L(\boldsymbol{x}, \lambda)$$

by setting the gradient of L with respect to $\boldsymbol{x}$ equal to zero (convex unconstrained problem, function in C^1). $\nabla_x L(\boldsymbol{x}, \lambda) = \mathbf{0} \Rightarrow x_i^* = -\frac{\lambda}{2}, \ \forall i$. We get $q(\lambda) = -\lambda b - n\frac{\lambda^2}{4}$.

In the Lagrangian dual problem we wish to maximize $q(\lambda)$ over $\mathbb{R}$ (no sign restrictions since the multiplier corresponds to an equality constraint). Also here, q is differentiable and we set the gradient equal to zero $\Rightarrow \lambda^* = -\frac{2b}{n}$ (we know that this is a maximum, since q is always concave) $\Rightarrow x_i^* = \frac{b}{n}, \ \forall i$.

Thus, for any faesible vector $\boldsymbol{x}$,

$$z^* = \sum_i \left(\frac{b}{n}\right)^2 = \frac{b^2}{n} \leq \sum_i x_i^2 \Leftrightarrow b^2 \leq n \sum_i x_i^2 \Leftrightarrow \left(\sum_{i=1}^{n} x_i\right)^2 \leq n \sum_i x_i^2.$$

The objective function is strictly convex, whence the inequality above holds with equality iff $x_i^* = \frac{b}{n}$, $\forall i$, i.e., if $x_1 = x_2 = \ldots = x_n$.
