Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 05–03–14
Examiner: Michael Patriksson

Question 1

(the Simplex method and sensitivity analysis in linear programming)

The problem in standard form is to

$$\begin{array}{lrrrrrl} \text{minimize} & z = -2x_1 & +(5+c)x_2 & -2x_3 & & & \\ \text{subject to} & x_1 & -3x_2 & +4x_3 & +x_4 & & = 2, \\ & 3x_1 & -x_2 & -3x_3 & & +x_5 & = 3-b, \\ & x_1, & x_2, & x_3 & & \geq & 0. \end{array}$$

(1p) a) The reduced costs of $\boldsymbol{x}_N = (x_2, x_4, x_5)^{\mathrm{T}}$ are $(2.2, 0.8, 0.4)^{\mathrm{T}} > (0,0,0)^{\mathrm{T}}$ which means that $\boldsymbol{x}_B = (x_1, x_3)^{\mathrm{T}}$ corresponds to the unique optimal solution.

(1p) b) For $b = 0$ the current basis is optimal if and only if $c \geq -11/5$, and for $c = 0$ the basis is optimal if and only if $-3 \leq b \leq 18/4$.

(1p) c) By choosing the entering and leaving variables according to the dual simplex method we get that x_3 is the leaving variable and x_2 the entering. The new basis becomes $\boldsymbol{x}_B = (x_1, x_2)^{\mathrm{T}}$, and it turns out that it is primal feasible and hence corresponds to an optimal solution to the modified problem.

(3p)

Question 2

(Newton's method)

We have

$$\nabla f(x,y) = (x^2/2, y)^{\mathrm{T}} \quad \text{and} \quad \nabla^2 f(x,y) = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence the first search direction is computed by solving the system

$$\begin{pmatrix} x_0 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{p}_0 = \begin{pmatrix} -x_0^2/2 \\ -y_0 \end{pmatrix} \quad \Longleftrightarrow \quad \boldsymbol{p}_0 = \begin{pmatrix} -x_0/2 \\ -y_0 \end{pmatrix}.$$

Hence we get that $\boldsymbol{x}_1 = \boldsymbol{x}_0 + \boldsymbol{p}_0 = (x_0/2, 0)^{\mathrm{T}}$, and it follows that the assertion is true for $k = 1$. Then use induction to show the general assertion.

The method converges to $(0,0)^{\mathrm{T}}$, which is not an optimal solution since the problem is unbounded.

Question 3

(Farkas' Lemma and other theorems of the alternative)

(2p) a) Farkas' Lemma is proved in the course notes.

(1p) b) We rewrite the system (I') by adding slack variables, thus producing

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} - \boldsymbol{I}^n \boldsymbol{s} &= \boldsymbol{b}, \\ (\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{s}^{\mathrm{T}})^{\mathrm{T}} &\geq \boldsymbol{0}^n \times \boldsymbol{0}^m. \end{aligned} \tag{J'}$$

This system is of the form (I) where the matrix $\boldsymbol{A}$ is replaced by $(\boldsymbol{A}, \boldsymbol{I}^n)$ and $\boldsymbol{x}$ by $(\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{s}^{\mathrm{T}})^{\mathrm{T}}$. Thus, we can apply Farkas' Lemma to this system and obtain a corresponding dual system,

$$\begin{aligned} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} &\leq \boldsymbol{0}^n, \\ -\boldsymbol{y} &\leq \boldsymbol{0}^n, \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} &> 0. \end{aligned} \tag{JJ'}$$

This system then has a solution if and only if (J') does not, and vice versa. Since the system (JJ') is the same as (II'), we have completed the proof.

Question 4

(optimality)

(1p) a) Abadie's CQ is fulfilled, since the four constraints all are linear (or, affine).

At $\boldsymbol{x}^*$ we satisfy all primal constraints, so it is *primal feasible*. The active constraints have the form

$$\begin{aligned} g_1(\boldsymbol{x}) &:= -x_1 + 1 \leq 0, \\ g_3(\boldsymbol{x}) &:= -x_2 + 1 \leq 0. \end{aligned}$$

Since we have that $\nabla f(\boldsymbol{x}^*) = (\mathrm{e}, 1)^{\mathrm{T}}$, solving the system of equations

$$\nabla f(\boldsymbol{x}^*) + \sum_{i \in \mathcal{I}(x^*)} \mu_i^* \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}^n$$

yields $\boldsymbol{\mu}^* = (\mathrm{e}, 0, 1, 0)^{\mathrm{T}}$. Since $\boldsymbol{\mu}^* \geq \boldsymbol{0}^4$, we satisfy the KKT conditions.

(1p) b) At $\boldsymbol{x}^*$ the matrix $\nabla^2 f(\boldsymbol{x}^*)$ is not positive semi-definite; it is actually indefinite, so the problem is not convex. Therefore, the fact that the KKT conditions are satisfied cannot be used to conclude that $\boldsymbol{x}^*$ is a global optimum.

However, we can conclude that it is a global optimum, in fact the unique global optimum, by studying the objective function on the feasible region. It is clear that the first term is non-negative on this set, and the second term is strictly increasing in x_1 and therefore has a minimum at $x_1 = 1$. A lower bound for the objective value on the feasible set therefore is e, which is exactly what is attained at $\boldsymbol{x}^*$. Hence, it is globally optimal.

(1p) c) Since the problem is convex, the KKT conditions imply that $\boldsymbol{x}^*$ is globally optimal, regardless of any CQ being fulfilled or not.

(3p) Question 5

(the variational inequality)

Consider the equivalent problem (in the sense that it has the same set of optimal solutions as the original problem) to

$$\begin{aligned} \text{minimize} \quad & g(\boldsymbol{x}) := -\ln\left(\sum_{i=1}^n c_i x_i\right) - \ln\left(\sum_{i=1}^n \frac{1}{c_i} x_i\right) \\ \text{subject to} \quad & \sum_{i=1}^n x_i = 1, \\ & x_i \geq 0, \quad i = 1, \dots, n. \end{aligned}$$

From the variational inequality it follows that

$$x_i > 0 \implies \frac{\partial g}{\partial x_i}(\boldsymbol{x}) \leq \frac{\partial g}{\partial x_j}(\boldsymbol{x}), \quad j = 1, \dots, n, \tag{1}$$

for every optimal solution $\boldsymbol{x}$. We have

$$\frac{\partial g}{\partial x_i}(\boldsymbol{x}) = -\frac{c_i}{\sum_{k=1}^n c_k x_k} - \frac{1}{c_i\left(\sum_{k=1}^n x_k / c_k\right)}.$$

Let $a := \sum_{k=1}^n c_k x_k$ and $b := \sum_{k=1}^n x_k / c_k$, and consider the function

$$h(t) := -at - \frac{1}{bt} \implies h(c_i) = \frac{\partial g}{\partial x_i}(\boldsymbol{x}).$$

We have that

$$h'(t) = -a + \frac{1}{bt^2}, \quad h''(t) = -\frac{2}{bt^3},$$

which means that h is strictly concave for all $t > 0$. Hence, since $c_1 < c_i < c_n$ for $i = 2, \dots, n-1$, it holds that

$$h(c_i) > \min\{h(c_1), h(c_n)\}, \quad i = 2, \dots, n-1.$$

This together with (1) imply that $x_2 = x_3 = \cdots = x_{n-1} = 0$ for every optimal solution. Now, assume that $x_1, x_n > 0$. Then by (1) it must hold that

$$h(c_1) = h(c_n) \quad \Longleftrightarrow \quad x_1 = x_n$$

and since $\sum_{i=1}^n x_i = 1$ we get $x_1 = x_n = 1/2$. The other possibilities are that $x_1 = 1$, $x_n = 0$, or $x_1 = 0$, $x_n = 1$. Assume that $x_1 = 1$ and $x_n = 0$. Then it follows that $h(c_1) = -2$. But we also have that

$$h(c_n) = -\frac{c_n}{c_1} - \frac{c_1}{c_n} = -\frac{(c_n - c_1)^2}{c_1 c_n} - 2 < -2,$$

which means that $h(c_1) > h(c_n)$, so (1) is not fulfilled and $x_1 = 1$, $x_n = 0$ cannot be an optimal solution. Similarly it follows that $x_1 = 0$, $x_n = 1$ cannot be optimal. Therefore we only have one solution, i.e. $x_1 = x_n = 1/2$, that might fulfill the variational inequality, and since the existence of an optimal solution is clear, this solution must be the unique optimal solution.

(3p) Question 6

(modelling)

Introduce the variables

$$x_{ijd} = \begin{cases} 1 & \text{if player } i \text{ meets player } j \text{ day } d, \\ 0 & \text{otherwise,} \end{cases}$$

$$z_{ij} = \begin{cases} 1 & \text{if there has been a meeting between player } i \text{ and } j \text{ during the week,} \\ 0 & \text{otherwise,} \end{cases}$$

and introduce the set $I = \{1, \dots, 20\}$. Then the following integer linear program solves the problem:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{19} \sum_{j=i+1}^{20} z_{ij} \\
\text{subject to} \quad & \sum_{j \in I \setminus \{i\}} x_{ijd} = 3, \quad i \in I, \quad d = 1, \dots, 7, \\
& x_{ijd} = x_{jid}, \quad i, j \in I, \quad d = 1, \dots, 7, \\
& x_{ikd} - x_{jkd} \leq 1 - x_{ijd}, \quad i \in I, \quad j \in I \setminus \{i\}, \quad k \in I \setminus \{i, j\}, \\
& z_{ij} \leq \sum_{d=1}^{7} x_{ijd}, \quad i = 1, \dots, 19, \quad j = i+1, \dots, 20, \\
& x_{ijd}, \; z_{ij} \in \{0, 1\}.
\end{aligned}$$

Note that the integer requirements on the z_{ij}-variables can be relaxed.

Question 7

(convex analysis)

(2p) a) We establish the result thus: $1 \implies 2 \implies 3 \implies 1$:

[$1 \implies 2$] By the statement 1., we have that $f(\boldsymbol{y}) \geq f(\boldsymbol{x}^*)$ for every $y \in \mathbb{R}^n$. This implies that for $\boldsymbol{g} = \boldsymbol{0}^n$, we satisfy the subgradient inequality (1). This establishes the statement 2.

[$2 \implies 3$] With $\boldsymbol{g} = \boldsymbol{0}^n$ the definition of $\partial f(\boldsymbol{x}^*)$ in (4) yields immediately the statement 3.

[$3 \implies 1$] By (3) and the compactness of the subdifferential (cf. Weierstrass' Theorem) the maximum is attained at some $\boldsymbol{g} \in \partial f(\boldsymbol{x}^*)$. It follows that, in the subgradient inequality, we get that

$$f(\boldsymbol{x}^* + \boldsymbol{p}) \geq f(\boldsymbol{x}^*) + \boldsymbol{g}^{\mathrm{T}} \boldsymbol{p} \geq f(\boldsymbol{x}^*), \qquad \forall p \in \mathbb{R}^n,$$

which is equivalent to the statement 1.

(1p) b) The answer is no.

Example 1: $f(x) := x^3$, and $x^* = 0$. This is an example where the derivative is zero, yet $p = -1$ is a descent direction.

Example 2: Any non-convex function $f \in C^2$ where $\boldsymbol{x}^*$ is a saddle point. In this case, $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}^n$, but there exists a descent direction given by an eigenvector corresponding to a negative eigenvalue of $\nabla^2 f(\boldsymbol{x}^*)$. Suppose that λ is a negative eigenvalue of $\nabla^2 f(\boldsymbol{x}^*)$, and that $\boldsymbol{p}$ is a corresponding eigenvector. Then,

$$\begin{aligned}
\nabla^2 f(\boldsymbol{x}^*)\boldsymbol{p} &= \lambda \boldsymbol{p} \qquad \Longrightarrow \\
\boldsymbol{p}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}^*)\boldsymbol{p} &= \lambda \|\boldsymbol{p}\|^2 < 0 \qquad \Longrightarrow \\
f(\boldsymbol{x}^* + \alpha \boldsymbol{p}) &= f(\boldsymbol{x}^*) + \alpha \nabla f(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{p} + \frac{\alpha^2}{2}\boldsymbol{p}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}^*)\boldsymbol{p} + o(\alpha^2) \\
&< f(\boldsymbol{x}^*)
\end{aligned}$$

for every small enough $\alpha > 0$.