Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 04–03–08

Examiner: Michael Patriksson

Question 1

(the Simplex method)

a) By introducing slack variables we get the problem in standard form:

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

The Phase I problem becomes

$$\begin{array}{lrrrrrrl}
\text{minimize} \quad w = & & & & & & a & \\
\text{subject to} & -2x_1 & +5x_2 & -x_3 & -x_4 & & +a & = 5, \\
& 2x_1 & -x_2 & +2x_3 & & +x_5 & & = 4, \\
& x_1, & x_2, & x_3, & x_4, & x_5, & a & \geq 0.
\end{array}$$

Start with the basis defined by $\boldsymbol{x}_B = (a, x_5)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_1, x_2, x_3, x_4)^{\mathrm{T}}$. The reduced costs of $\boldsymbol{x}_N$ become $(2, -5, 1, 1)$, so x_2 is the entering variable. The leaving variable becomes a. The new basis is given by $\boldsymbol{x}_B = (x_2, x_5)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_1, a, x_3, x_4)^{\mathrm{T}}$, and the reduced costs of $\boldsymbol{x}_N$ are $(0, 1, 0, 0)$, which means that the current basis is optimal to the Phase I problem and since $w^* = 0$ it follows that $\boldsymbol{x}_B = (x_2, x_5)^{\mathrm{T}}$, $\boldsymbol{x}_N = (x_1, x_3, x_4)^{\mathrm{T}}$ define a BFS to the Phase II problem (P). The reduced costs of $\boldsymbol{x}_N$ becomes $(2.2, 1.6, 0.6)^{\mathrm{T}} \geq \mathbf{0}^3$, which means that an optimal solution to (P) is given by

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} x_2 \\ x_5 \\ x_1 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence an optimal solution to the original problem is given by

$$\boldsymbol{x}^* = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

b) Since the reduced costs of $\boldsymbol{x}_N$ are all strictly positive, it follows that the BFS found is the unique optimal solution (see Proposition 10.9 in the course notes).

Question 2

(optimality conditions)

a) Drawing the figure one can verify that the problem is non-convex, because the feasible set is not convex (even though the objective function is). The optimization problem amounts to finding the shortest distance from the point $(x, y)^{\mathrm{T}} = (2, 1)^{\mathrm{T}}$ to the feasible set, and the geometrical considerations give us one local minimum $(x, y)^{\mathrm{T}} = (2, 0)^{\mathrm{T}}$ with the objective value $f((2, 0)^{\mathrm{T}}) = 1/2$ and a global minimum $(x, y)^{\mathrm{T}} = (3/2, 3/2)^{\mathrm{T}}$ with objective value $f((3/2, 3/2)^{\mathrm{T}}) = 1/4$.

Introducing the KKT-multipliers μ_1 and μ_2 for the inequality constraints, as well as λ for the equality constraint, the KKT system for this problem can be stated as follows:

$$\begin{cases} \begin{pmatrix} x-2 \\ y-1 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} \mu_1 + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \mu_2 + \begin{pmatrix} y \\ x-2y \end{pmatrix} \lambda = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ y - x \leq 0, \\ -y \leq 0, \\ y(x-y) = 0, \\ \mu_1, \mu_2 \geq 0, \\ \mu_1(x-y) = 0, \\ \mu_2 y = 0. \end{cases}$$

As it can be verified, this system gives two [in the space $(x, y)^{\mathrm{T}}$] KKT-points:

- The point of local minimum: $(x, y)^{\mathrm{T}} = (2, 0)^{\mathrm{T}}$, $\mu_1 = 0$, $\mu_2 \geq 0$, $2\lambda = 1 + \mu_2$.
- The point of global minimum: $(x, y)^{\mathrm{T}} = (3/2, 3/2)^{\mathrm{T}}$, $\mu_1 \geq 0$, $\mu_2 = 0$, $3\lambda = 1 + 2\mu_1$.

b) A simple calculation shows that the gradients of the free constraints are: $\nabla g_1(x, y) = (1, -1)^{\mathrm{T}}$, $\nabla g_2(x, y) = (0, 1)^{\mathrm{T}}$, $\nabla g_3(x, y) = (y, x - 2y)^{\mathrm{T}}$. At every feasible point we have either $y = 0$, which results in $\nabla g_2(x, y) = x\nabla g_3(x, y)$, or $x = y$, which results in $\nabla g_1(x, y) = y\nabla g_3(x, y)$. In either case, the LICQ is violated.

Again, from either geometrical or analytical considerations, we can split the feasible set of the original problem into two (non-disjoint) parts defined by linear constraints:

$$\mathcal{F}_1 = \{ (x, y) \in \mathbb{R}^2 \mid y = 0, x - y \geq 0 \},$$

and

$$\mathcal{F}_2 = \{ (x, y) \in \mathbb{R}^2 \mid y \geq 0, x - y = 0 \}.$$

We can therefore solve two convex linearly constrained optimization problems:

$$\begin{aligned} &\text{minimize} \;\; f(x, y), \\ &\text{subject to} \;\; (x, y) \in \mathcal{F}_1, \end{aligned}$$

and

$$\begin{aligned} &\text{minimize} \;\; f(x, y), \\ &\text{subject to} \;\; (x, y) \in \mathcal{F}_2, \end{aligned}$$

and choose the best solution among the two.

c) The procedure in the previous part can be generalized for problems with several complementarity constraints as follows. The feasible set can be split into 2^n parts $\mathcal{F}_I$, $I \subseteq \{1, \ldots, n\}$, where

$$\begin{aligned} \boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{x} = b_i, \text{and } x_i \geq 0, \quad i \in I, \\ \boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{x} \geq b_i, \text{and } x_i = 0, \quad i \notin I. \end{aligned}$$

Therefore, instead of solving the origial non-convex problem, which violates the LICQ, one can (in principle) solve 2^n convex problems with linear constraints.

Question 3

(modelling)

Introduce variables according to Figure 1.

Introduce constraints according to the following list:

Maximum sales:

$$x_1 \leq 200, \quad x_2 \leq 100, \quad x_3 \leq 300. \tag{1}$$

Process balances, Machine 1:

$$y_1 \geq x_1, \quad y_2 \geq x_2, \quad y_3 \geq x_2, \quad y_4 \geq x_3. \tag{2}$$

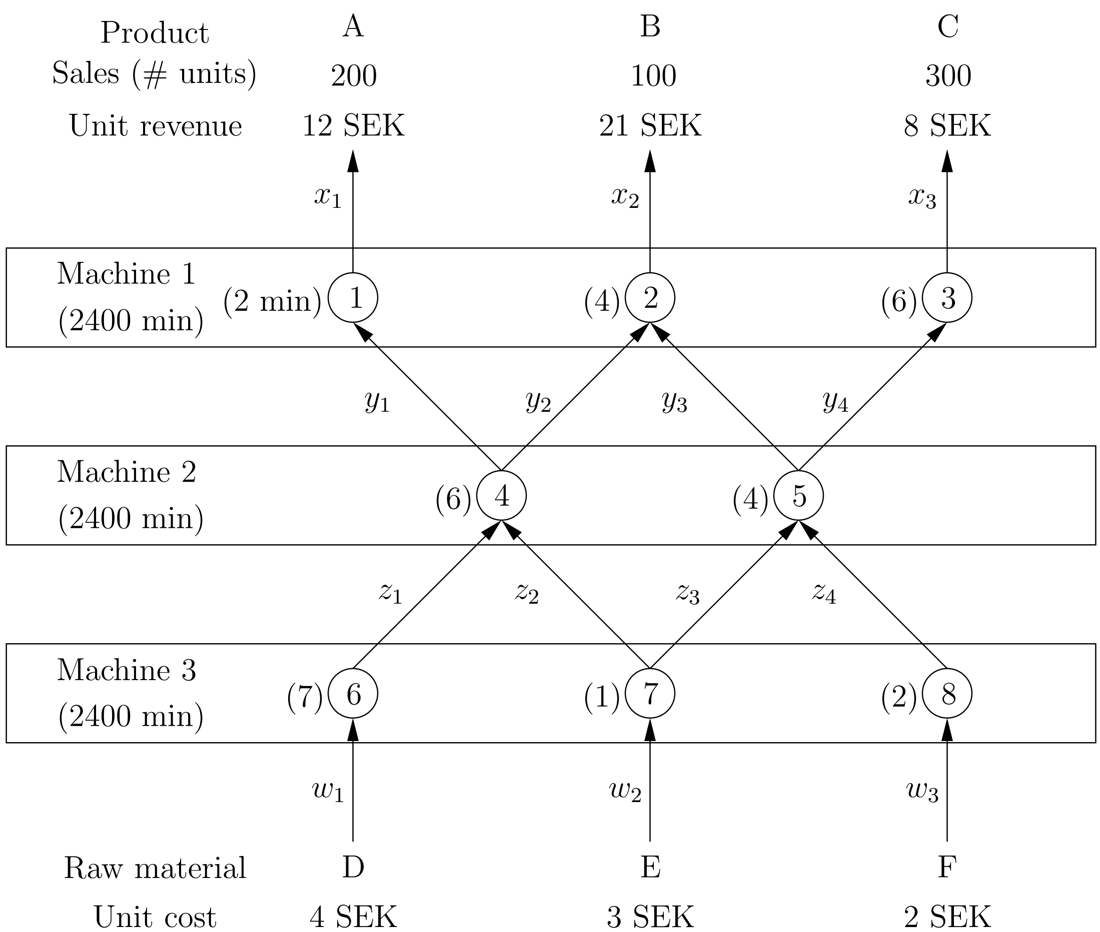

Figure 1: Variable definitions.

Process balances, Machine 2:

$$z_1 \geq y_1 + y_2, \quad z_2 \geq y_1 + y_2, \quad z_3 \geq y_3 + y_4, \quad z_4 \geq y_3 + y_4. \tag{3}$$

Process balances, Machine 3:

$$w_1 \geq z_1, \quad w_2 \geq z_2 + z_3, \quad w_3 \geq z_4. \tag{4}$$

Weekly capacity, Machine 1:

$$2x_1 + 4x_2 + 6x_3 \leq 2400. \tag{5}$$

Weekly capacity, Machine 2:

$$6(y_1 + y_2) + 4(y_3 + y_4) \leq 2400. \tag{6}$$

Weekly capacity, Machine 3:

$$7z_1 + (z_2 + z_3) + 2z_4 \leq 2400. \tag{7}$$

Objective function:

$$f(\boldsymbol{x}, \boldsymbol{w}) = 12x_1 + 21x_2 + 8x_3 - 4w_1 - 3w_2 - 2w_3.$$

We end up with the linear integer program

$$\begin{aligned} \text{maximize} \quad & f(\boldsymbol{x}, \boldsymbol{w}), \\ \text{subject to} \quad & (1) - (7), \\ & \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{w} \geq \mathbf{0} \text{ and integer.} \end{aligned}$$

Question 4

(applications of the Newton algorithm)

a) The objective function $f(x) = ax - \log(x)$ is strictly convex inside the feasible set $\{\, x \in \mathbb{R} \mid x > 0 \,\}$, since $f''(x) = 1/x^2 > 0$ there; therefore, every local minimum in this problem is also a global one, and the global minimum is unique, provided any exists. Now we can test the necessary (and sufficient in this case, owing to the convexity) optimality conditions

$$\begin{aligned} f'(x) = a - x^{-1} &= 0, \\ x &> 0, \end{aligned}$$

which is uniquely solvable, giving us $x^* = a^{-1} > 0$.

b) Direct calculations show that

$$x_{k+1} = x_k - f'(x_k)/f''(x_k) = x_k(2 - ax_k),$$

which does not involve any divisions.

Assuming that $x_k \to \bar{x}$ (and thus also $x_{k+1} \to \bar{x}$) gives us

$$\bar{x} = \bar{x}(2 - a\bar{x}),$$

which has two solutions: $\bar{x}_1 = a^{-1}$ or $\bar{x}_2 = 0$. It is the latter solution that is not a global/local optimum of the original problem (it is not even feasible, to start with). One can easily obtain this solution by starting from the point $x_0 = 2/a > 0$, which generates $x_1 = 0$, and thus $x_k = 0$ for all $k \geq 1$.

c) One can for example start from the optimality conditions

$$g'(x) = a - x^{-2} = 0,$$
$$x > 0,$$

to end up with the strictly convex minimization problem to

$$\begin{aligned} &\text{minimize} \;\; g(x) = ax + x^{-1}, \\ &\text{subject to} \;\; x > 0. \end{aligned}$$

It is verified as in b) that Newton's method for this problem involves only simple operations (additions/subtractions and multiplications).

Question 5

(optimality conditions)

Farkas' Lemma can be stated as follows:

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{b}$ an $m \times 1$ vector. Then exactly one of the systems

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b}, \\ \boldsymbol{x} &\geq \boldsymbol{0}^n, \end{aligned} \tag{I}$$

and

$$\begin{aligned} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} &\leq \boldsymbol{0}^n, \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} &> 0, \end{aligned} \tag{II}$$

has a feasible solution, and the other system is inconsistent.

a) Farkas' Lemma is proved in Theorem 11.10.

b) At $\bar{\boldsymbol{x}} := (0,0)^{\mathrm{T}}$, the cone of feasible directions is

$$\begin{aligned} R_S(\bar{\boldsymbol{x}}) &= \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid 2p_1 - p_2 = 0;\; \boldsymbol{p} \geq \boldsymbol{0}^2 \,\} \\ &= \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid 2p_1 - p_2 \leq 0;\; -2p_1 + p_2 \leq 0;\; -p_1 \leq 0;\; -p_2 \leq 0 \,\}. \end{aligned}$$

At $\bar{\boldsymbol{x}} := (0,0)^{\mathrm{T}}$, the cone of descent directions is

$$\overset{\circ}{F}(\bar{\boldsymbol{x}}) = \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{p} < 0 \,\} = \{\, \boldsymbol{p} \in \mathbb{R}^2 \mid p_1 + p_2 > 0 \,\}.$$

To prove that the set $R_S(\bar{\boldsymbol{x}}) \cap \overset{\circ}{F}(\bar{\boldsymbol{x}})$ is non-empty (that is, that there exists a feasible descent direction), we define

$$\boldsymbol{A} := \begin{pmatrix} 2 & -2 & -1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The consistency of the system (II) then is equivalent to the existence of a feasible descent direction (with $\boldsymbol{p} = \boldsymbol{y}$). We therefore need to establish that the system (I) is inconsistent. The consistency of this system is equivalent to the possibility to choose a non-negative $\boldsymbol{x} \in \mathbb{R}^4$ such that

$$\begin{pmatrix} 2 & -2 & -1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

This is however impossible. (One way to check this is via Phase I in the Simplex method.)

We are done.

Question 6

(convexity)

The proof of Carathéodory's Theorem can be found in Theorem 3.8 in the Course Notes.

Question 7

(duality in linear and nonlinear optimization)

a) The LP dual is to

$$\begin{array}{lllll} \text{maximize} & w = \boldsymbol{b}_1^{\mathrm{T}} \boldsymbol{y}_1 & + \boldsymbol{b}_2^{\mathrm{T}} \boldsymbol{y}_2 & + a y_3 & \\ \text{subject to} & \boldsymbol{A}_1^{\mathrm{T}} \boldsymbol{y}_1 & + \boldsymbol{A}_2^{\mathrm{T}} \boldsymbol{y}_2 & & \leq \boldsymbol{c}, \\ & \boldsymbol{B}^{\mathrm{T}} \boldsymbol{y}_1 & & + \boldsymbol{1}^{\ell} y_3 & \leq \boldsymbol{d}, \\ & \boldsymbol{y}_1 \geq \boldsymbol{0}^{m_1}, & \boldsymbol{y}_2 \in \mathbb{R}^{m_2}, y_3 \in \mathbb{R}, & & \end{array}$$

where $\boldsymbol{1}^{m_1}$ is the m_1-vector of ones.

b) With $g(\boldsymbol{x}) := -x_1 + 2x_2 - 4$, the Lagrange function becomes

$$\begin{aligned} L(\boldsymbol{x}, \mu) &= f(\boldsymbol{x}) + \mu g(\boldsymbol{x}) \\ &= 2x_1^2 + x_2^2 - 4x_1 - 6x_2 + \mu(-x_1 + 2x_2 - 4). \end{aligned}$$

Minimizing this function over $\boldsymbol{x} \in \mathbb{R}^2$ yields [since $L(\cdot, \mu)$ is a strictly convex quadratic function for every value of μ, it has a unique minimum for every value of μ] that its minimum is attained where its gradient is zero. This gives us that

$$\begin{aligned} x_1(\mu) &= (4 + \mu)/4; \\ x_2(\mu) &= 3 - \mu. \end{aligned}$$

Inserting this into the Lagrangian function, we define the dual objective function as

$$q(\mu) = L(\boldsymbol{x}(\mu), \mu) = \cdots = -2\left(\frac{4 + \mu}{4}\right)^2 - (3 - \mu)^2 - 4\mu.$$

This function is to be maximized over $\mu \geq 0$. We are done with task [1].

We attempt to optimize the one-dimensional function q by setting the derivative of q to zero. If the resulting value of μ is non-negative, then it must be a global optimum; otherwise, the optimum is $\mu^* = 0$.

We have that $q'(\mu) = \cdots = 1 - \frac{9\mu}{4}$, so the stationary point of q is $\mu = 4/9$. Since its value is positive, we know that the global maximum of q over $\mu \geq 0$ is $\mu^* = 4/9$. We are done with task [2].

Our candidate for the global optimum in the primal problem is $\boldsymbol{x}(\mu^*) = \frac{1}{9}(10, 23)^{\mathrm{T}}$. Checking feasibility, we see that $g(\boldsymbol{x}(\mu^*)) = 0$. Hence, without even evaluating the values of $q(\mu^*)$ and $f(\boldsymbol{x}(\mu^*))$ we know they must be equal, since $q(\mu^*) = f(\boldsymbol{x}(\mu^*)) + \mu^* g(\boldsymbol{x}(\mu^*)) = f(\boldsymbol{x}(\mu^*))$, due to the fact that we satisfy complementarity. We have proved that strong duality holds, and therefore task [4] is done.

By the Weak Duality Theorem 7.4 follows that if a vector $\boldsymbol{x}$ is primal feasible and $f(\boldsymbol{x}) = q(\mu)$ holds for some feasible dual vector μ, then $\boldsymbol{x}$ must be the optimal solution to the primal problem. (And μ must be optimal in the dual problem.) Task [4] is completed by the remark that this is exactly the case for the pair $(\boldsymbol{x}(\mu^*), \mu^*)$.