Tentamen / Exam TMA946 / MAN280 03-03-10

SOLUTION OF QUESTION 1

NICLAS ANDRÉASSON

By introducing a slack variable x_5 and two artificial variables a_1 and a_2 we get the Phase I problem

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

Let $x_B = [a_1, a_2, x_5]$ and $x_N = [x_1, x_2, x_3, x_4]$ be the initial basic and nonbasic vector respectively. The reduced costs of the nonbasic variables then become

$$c_N^T - c_B^T B^{-1} N = [-2, 1, 1, 2],$$

and thus x_1 is the entering variable. Further, we have

$$B^{-1}b = [3, 1, 7]^T,$$
$$B^{-1}N_1 = [1, 1, 2]^T,$$

which gives

$$\underset{j,\,(B^{-1}N_1)_j > 0}{\arg\min} \frac{B^{-1}b}{(B^{-1}N_1)_j} = 2,$$

so a_2 is the leaving variable. The new basic and nonbasic vectors are $x_B = [a_1, x_1, x_5]$ and $x_N = [a_2, x_2, x_3, x_4]$, and the reduced costs of the nonbasic variables become

$$c_N^T - c_B^T B^{-1} N = [2, -1, 1, -2],$$

so x_4 is the entering variable, and

$$B^{-1}b = [2, 1, 5]^T,$$
$$B^{-1}N_4 = [2, -2, 5]^T,$$

which gives

$$\underset{j,\,(B^{-1}N_4)_j > 0}{\arg\min} \frac{B^{-1}b}{(B^{-1}N_4)_j} = 1,$$

and thus a_1 is the leaving variable. The new basic and nonbasic vectors become $x_B = [x_4, x_1, x_5]$ and $x_N = [a_2, x_2, x_3, a_1]$, and the reduced costs of the nonbasic variables are

$$c_N^T - c_B^T B^{-1} N = [1, 0, 0, 1],$$

so $x_B = [x_4, x_1, x_5]$ is an optimal basic feasible solution of the Phase I problem, and $w^* = 0$. This means that $x_B = [x_4, x_1, x_5]$ is a basic feasible solution of the Phase II problem, i.e.,

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

If $x_B = [x_4, x_1, x_5]$ and $x_N = [x_2, x_3]$ we get the reduced costs

$$c_N^T - c_B^T B^{-1} N = [0, 2],$$

which means that $x_B = [x_4, x_1, x_5]$ is an optimal basic feasible solution. But since the reduced cost of x_2 is zero there is a possibility that there are alternative solutions. Let x_2 enter the basic vector. Then we get

$$B^{-1}b = [1, 3, 0]^T,$$
$$B^{-1}N_1 = [0.5, 0, -0.5]^T,$$

which gives

$$\underset{j,(B^{-1}N_1)_j>0}{\arg\min} \frac{B^{-1}b}{(B^{-1}N_1)_j} = 1,$$

so x_4 is the leaving variable. We get $x_B = [x_2, x_1, x_5]$ and $x_N = [x_4, x_3]$, and the reduced costs become

$$c_N^T - c_B^T B^{-1} N = [0, 2],$$

so $x_B = [x_2, x_1, x_5]$ is an optimal basic feasible solution, and since

$$B^{-1}b = [2, 3, 1]^T$$

it is an alternative optimal solution.

Answer: The solution of the given LP is not unique. Two different optimal solutions are given by

$$x_1^* = [x_1, x_2, x_3, x_4] = [3, 0, 0, 1],$$
$$x_2^* = [x_1, x_2, x_3, x_4] = [3, 2, 0, 0],$$

and

$$z^* = 6.$$

1 Question 2

a)

If $x^*(b)$ is a basic feasible solution with $B^{-1}b > 0$ we have that the basis stays feasible for small changes in b. The optimality of the basis is not affected by a change in b. As we have that the optimal value is $z = c_b^{\mathrm{T}} B^{-1} b$ we get $\nabla_b z(b) = (c_b^{\mathrm{T}} B^{-1})^{\mathrm{T}}$.

b)

Take two right hand sides b_1 and b_2 with the corresponding optimal solutions x_1 and x_2. We now form $b(\alpha)$ as $b(\alpha) = \alpha b_1 + (1-\alpha) b_2, \alpha \in [0,1]$ and $x(\alpha)$ as $x(\alpha) = \alpha x_1 + (1-\alpha)x_2, \alpha \in [0,1]$.

We have that

$$Ax(\alpha) = A(\alpha x_1 + (1-\alpha)x_2) = \alpha A x_1 + (1-\alpha)Ax_2 = \alpha b_1 + (1-\alpha)b_2 = b(\alpha)$$

Hence $x(\alpha)$ is a feasible solution to the problem

$$\begin{aligned} \min \quad & c^{\mathrm{T}} x && (1)\\ \text{subject to} \quad & Ax = b(\alpha) && (2)\\ & x \geq 0 && (3)\\ & && (4) \end{aligned}$$

As $x(\alpha)$ is feasible but not necessarily optimal, we get that

$$z^*(b(\alpha)) \leq c^{\mathrm{T}} x(\alpha) = \alpha c^{\mathrm{T}} x_1 + (1-\alpha) c^{\mathrm{T}} x_2 = \alpha z^*(b_1) + (1-\alpha) z^*(b_2)$$

We now have shown that $z^*(b)$ is a convex function.

2 Question 3

We define the following variables.

x_c	Amount of clay picked from pit.
x_l	Amount of limestone bought from quarry.
y_c	Amount of clay after drying.
y_l	Amount of limestone after crushing.
v	Amount of cement produced.
w_i	Amount of cement sold to site i.
p	the price we set.
M_i	Indicator variable, 1 if we sell to site i, 0 otherwise.
H	A sufficiently huge value.

The objective is to minimise profits from sales, after costs.

$$\text{maximise} \sum_{i=1}^{3} w_i p - q x_l - c x_c - d x_l$$

The constraints are,

$x_l + x_c \leq b$	We must not exceed maximum transport capacity
$x_l = y_l$	The weight of crushed limestone equals weight of input
$0.7x_c = y_c$	Clay looses 30% of weight when drying
$v = y_l + y_c$	Cement is clay and limestome
$1.7y_l \leq y_c$	minimum amount of clay required
$2.1y_l \geq y_c$	maximum amount of clay allowed
$\sum_{i=1}^{3} w_i = v$	We must produce what we sell
$w_i = M_i \frac{k_i}{m_i + p^2}, i = 1 \ldots, 3$	Sold amount depends on price
$p \leq (1 - M_i)H + r_i$	can't sell if price is too high

Furthermore, we have sign restrictions

$$x_c, x_l, y_c, y_l, v, w, p \geq 0$$

And integrality restrictions

$$M_i \in \{0, 1\}, i \in 1, \ldots, 3$$

4. a) The KKT conditions: for x^* being a local minimum of the problem, there exists a vector λ^* with

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*), \qquad (1)$$

$$\lambda^* \geq 0^m, \qquad (2)$$

$$\lambda_i^* \cdot g_i(x^*) = 0, \quad i = 1, \ldots, m \qquad (3)$$

$$g_i(x^*) \geq 0, \quad i = 1, \ldots, m \qquad (4)$$

Because of the convexity of the problem, the Lagrangian function $L(\cdot, \lambda^*)$ is convex. It is also continuously differentiable, by the assumptions. Therefore, the set $(X = \mathbb{R}^n)$ $\{x \in X \mid L(x, \lambda^*) = L_*(\lambda^*)\}$ equals $\{x \in \mathbb{R}^n \mid \nabla_x L(x, \lambda^*) = 0^n\}$, which is identical to the condition (1) above. The second set in the global optimality conditions corresponds to (4), and the third to (3). The condition (2) completes the set of global optimality conditions and KKT conditions; so, the two are equivalent.

b) The condition $L(x^*, \lambda^*) = L_*(\lambda^*)$ is identical to $f(x^*) = L_*(\lambda^*)$, thanks to the condition that $(\lambda^*)^T g(x^*) = 0$. By weak duality, $f(x^*) \geq L_*(\lambda^*)$ for every feasible solution in the primal and dual problems. It follows that x^* must be primal optimal and λ^* must be dual optimal.

c) The dual function has the form $\mathcal{L}_*(\lambda) = \underset{x \in X}{\text{minimum}}\ \mathcal{L}(x, \lambda^*)$
$= f(x(\lambda^*)) - (\lambda^*)^T g(x(\lambda^*))$ for some $x(\lambda^*) \in X$,
by Weierstrass' Theorem. Moreover, for any $\lambda \in \mathbb{R}^m$,

$$\mathcal{L}_*(\lambda) = \underset{x \in X}{\text{minimum}}\ \mathcal{L}(x, \lambda)$$

$$\le \mathcal{L}(x(\lambda^*), \lambda)$$
$$= \mathcal{L}(x(\lambda^*), \lambda^*) + [\mathcal{L}(x(\lambda^*), \lambda) - \mathcal{L}(x(\lambda^*), \lambda^*)]$$
$$= \mathcal{L}_*(\lambda^*) - g(x(\lambda^*))^T(\lambda - \lambda^*),$$

which shows that $-g(x(\lambda^*))$ is a subgradient of $\mathcal{L}_*$ at λ^*.
The function $\mathcal{L}_*$ is concave on $\mathbb{R}^m$. Therefore, if $\mathcal{L}_*$ is differentiable, it must follow that $\nabla \mathcal{L}_*(\lambda^*) = -g(x(\lambda^*))$.
Conversely, if $g(x(\lambda^*))$ is the same for every $x(\lambda^*)$ that satisfies $\mathcal{L}_*(\lambda^*) = \mathcal{L}(x(\lambda^*), \lambda^*)$, then the set of subgradients is a singleton, and $\nabla \mathcal{L}_*(\lambda^*)$ exists.

$\therefore$ $\nabla \mathcal{L}_*(\lambda^*)$ exists, and equals $-g(x(\lambda^*))$ for any vector $x(\lambda^*)$ with $\mathcal{L}_*(\lambda^*) = \mathcal{L}(x(\lambda^*), \lambda^*)$ if and only if $\{g(x(\lambda^*)) \mid \mathcal{L}_*(\lambda^*) = \mathcal{L}(x(\lambda^*), \lambda^*)\}$ is a singleton set.

A sufficient condition for the above to hold is that f is strictly convex on X while g_i ($i = 1, \ldots, m$) are concave, since then $x(\lambda)$ is unique for every $\lambda \in \mathbb{R}^m$.

The dual problem is of the form: $\underset{\lambda \ge 0^m}{\text{maximize}}\ \mathcal{L}_*(\lambda)$.
As $\mathcal{L}_*$ is differentiable at the optimal solution λ^* by assumption, the optimality conditions for this problem are that

$$\lambda^* \ge 0^m;\quad \nabla L_*(\lambda^*) \le 0^m;\quad (\lambda^*)^T \nabla L_*(\lambda^*) = 0.$$

This comes out to be: $\lambda^* \geq 0^m$; $g(x(\lambda^*)) \geq 0^m$; $(\lambda^*)^T g(x(\lambda^*)) = 0$.
Hence, under the differentiability assumption, <u>any</u> vector $x(\lambda^*)$ satisfying $\mathcal{L}_*(\lambda^*) = \mathcal{L}(x(\lambda^*), \lambda^*)$ is primal feasible and complementary to λ^*, hence optimal in the primal problem.
(Note that $\mathcal{L}_*(\lambda^*) = \mathcal{L}(x(\lambda^*), \lambda^*) \Leftrightarrow f(x(\lambda^*)) = \mathcal{L}_*(\lambda^*)$.)

5a) See the text book.

b) At $\bar{x}$, the incoming criterion in the simplex method is based on the problem to

minimize $w := \bar{c}_N^T p_N$, where $\bar{c}_N^T = c_N^T - c_B^T B^{-1} N$

subject to $p_N \geq 0$; $\sum_{j \in N} p_j \leq 1$.

If $\bar{c}_N^T \geq 0$, then $w^* = 0$ and we terminate; otherwise, $w^* < 0$ and we choose $p_{j^*} = 1$ for one $j^* \in N$, and $p_j = 0$, $j \in N \setminus \{j^*\}$. That the problem can be seen as a restriction becomes clear when we note that $p_N = x_N$ while $x_B \geq 0$ corresponds to $B^{-1}b - B^{-1}Np_N \geq 0$. If we associate p_B with $-B^{-1}Np_N$, we see (a) that p has at most $m+1$ nonzeros, and (b) that the addition of the constraints $p_B = -B^{-1}Np_N$ implies that p is a feasible direction: $Ap = Bp_B + Np_N = -Np_N + Np_N = 0$. If $\bar{x}$ is nondegenerate, p is also feasible with respect to the non-negativity constraints.

⑥ $\rho(x) := \frac{x^t A x}{x^t x}$

a) Stationarity $\Leftrightarrow \nabla \rho(x) = 0$

$$\nabla \rho(x) = \frac{2(x^t x \cdot Ax - x^t Ax \cdot x)}{(x^t x)^2} = \frac{2}{x^t x}(Ax - \rho(x)\cdot x) = 0 \quad (*)$$

If $x_i \neq 0$ is an eigenvector of A, corresponding to an eigenvalue $\lambda_i \Rightarrow \rho(x_i) = \frac{x^t Ax}{x^t x} = \lambda_i \frac{x^t x}{x^t x} = \lambda_i \quad (**)$

From $(*)$ & $(**)$ it follow that for $x \neq 0$ to be a stationary point of $\rho(\cdot)$, it is both necessary & sufficient for x to be an eigenvector.

The global minimum is an arbitrary nonzero eigenvector, corresponding to the minimal eigenvalue λ_{min}.

b) $(\hat{P})$ is non-convex iff $\nabla^2(x^t Ax) = 2A \prec 0$, i.e., $0 > \lambda_{min} = \rho^*$, the optimal value of (P).

KKT-conditions for the problem (P) [which are necessary for optimality] are:

$$\begin{cases} 2Ax + 2\mu x = 0 \\ \mu \geq 0, \quad x^t x \leq 1 \\ \mu^t(1 - x^t x) = 0 \end{cases}$$

Therefore, the KKT-points are:

1) $x \in \text{null}(A) \cap \{x^t x \leq 1\}$, with $\mu = 0$ & $x^t Ax = 0$

2) Eigenvectors of unit length, corresponding to negative [there is at least one, owing to non-convexity] eigenvalues of A, with $\mu = -\lambda_i > 0$, $x^t Ax = \lambda_i \|x\|^2 = \lambda_i$

Therefore, the optimal objective value is $\lambda_{min} < 0$!

7

$(P) \begin{cases} \min\ f(x) \\ \text{s.t. } g_i(x) \le 0 \end{cases}$ $(\hat{P}) \begin{cases} \min\ \hat{f}(y) \\ \text{s.t. } \hat{g}_i(y) \le 0 \end{cases}$

KKT-(P): $\begin{cases} \nabla f(x) + \sum_i \lambda_i \nabla g_i(x) = 0 \\ \lambda_i \geq 0,\ g_i(x) \le 0 \\ \lambda_i\, g_i(x) = 0 \end{cases}$

$y_j = \alpha_j x_j$ $\hat{f} = \beta_0 f$ $\hat{g}_i = \beta_i g$

$$\frac{\partial \hat{f}}{\partial y_j}(y) = \beta_0 \frac{\partial f}{\partial x_j}(x) \cdot \frac{\partial x_j}{\partial y_j} = \alpha_j^{-1} \cdot \beta_0 \cdot \frac{\partial f}{\partial x_j}(x)$$

$$\nabla \hat{f}(y) = \beta_0 \cdot \alpha^{-1} * \nabla f(x)$$

In the same way, $\nabla \hat{g}_i(y) = \beta_i \cdot \alpha^{-1} * \nabla g_i(x)$

Therefore, if x_0 satisfies KKT-(P), then $y_0 := \alpha * x_0$ satisfies:

$$\begin{cases} \beta_0^{-1} \cdot \alpha * \nabla \hat{f}(y_0) + \sum \lambda_i \beta_i^{-1} \cdot \alpha * \nabla \hat{g}_i(y_0) = 0 \\ \lambda_i \geq 0,\ \hat{g}_i(y_0) \le 0 \\ \lambda_i \cdot \hat{g}_i(y) = 0 \end{cases}$$

Defining $\hat{\lambda}_i := \lambda_i \cdot \frac{\beta_0}{\beta_i}$, this is equivalent

$$\begin{cases} \nabla \hat{f}(y_0) + \sum \hat{\lambda}_i \cdot \nabla \hat{g}_i(y_0) = 0 \\ \hat{\lambda}_i \geq 0,\ \hat{g}_i(y_0) \le 0 \\ \hat{\lambda}_i \cdot \hat{g}_i(y) = 0 \end{cases}$$

which is exactly the KKT-conditions for $(\tilde{P})$.

Therefore, $\hat{\lambda}_i = \beta_0 \cdot \beta^{-1} * \lambda$

7

b). While $f(x)$ is differentiable everywhere,

$\sqrt{f(x)}$ is not differentiable @ $x = c$ (equivalently, if $f(x) = 0$), which is the main advantage of (P) over $(\hat{P})$.

Therefore, if KKT-conditions are satisfied for (P), i.e.

$$\begin{cases} \nabla f(x) + \sum \lambda_i \nabla g_i(x) = 0 \\ \lambda_i \geq 0, \quad g_i(x) \leq 0 \\ \lambda_i \cdot g_i(x) = 0 \end{cases}$$

and $f(x) \neq 0$, then

$$\nabla \sqrt{f(x)} = \frac{\nabla f(x)}{2\sqrt{f(x)}}$$, and KKT for $(\hat{P})$

are satisfied with $\hat{\lambda}_i = \dfrac{\lambda_i}{2\sqrt{f(x)}}$.

Otherwise (if $f(x) = 0$), x is not a KKT-point for $(\hat{P})$ (Counter-example: no constraints, $n = 1$, $C = 0 \Rightarrow f(x) = x^2$; $\sqrt{f(x)} = |x|$, $x = 0$).

On the other hand, if x^* is a global optimum, i.e., $g_i(x^*) \leq 0$, $\forall x: g_i(x) \leq 0 \Rightarrow 0 \leq f(x^*) \leq f(x)$

it follows that $0 \leq \sqrt{f(x^*)} \leq \sqrt{f(x)}$, because $\sqrt{\cdot}$ is a monotonically increasing function.

I.e., x^* is also a global optimum in $(\hat{P})$