Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date: 12–08–30

Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) We first rewrite the problem on standard form. We multiply the objective and first constraint by -1 and introduce slack variables s_1 and s_2.

$$\begin{array}{lrrrrrl} \text{minimize} & z = & -x_1 & -2x_2 & & & \\ \text{subject to} & & -x_1 & +3x_2 & +s_1 & & = 6 \\ & & x_1 & +x_2 & & +s_2 & = 1 \\ & & x_1, & x_2, & s_1, & s_2 & \geq 0. \end{array}$$

By choosing (s_1, s_2) as basic variables, we obtain a BFS with $B = I$, hence we can begin with phase II.

Calculating the vector of reduced costs for the non-basic variables x_1, x_2 yields $(-1, -2)^{\mathrm{T}}$. Least reduced cost implies that x_2 is the entering variable. The minimum ratio test shows that s_1 should leave the basis. We have the basic variables (x_2, s_2). The vector of reduced costs for the non-basic variables x_1 and s_1 is $(-5/3, 2/3)^{\mathrm{T}}$. Hence x_1 enters the basis. The minimum ratio test implies that s_2 leaves the basis. We now have x_1, x_2 as basic variables. The vector of reduced costs for the non-basic variables s_1, s_2 is $(3/2, 5/2)^{\mathrm{T}}$. Since the reduced costs are all non-negative, the current BFS is optimal. We obtain $(x_1, x_2)^{\mathrm{T}} = B^{-1}b = (9/2, 7/2)^{\mathrm{T}}$ as the optimal solution and 23 as the optimal solution value for the original problem (-23 for the problem in standard form).

(1p) b) In the optimal BFS, all the reduced costs are strictly greater than zero. Hence the optimal solution is unique.

(3p)

Question 2

(Newton's method)

The search direction for Newton's method is defined by solving the linear system

$$\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{p}_{k+1} = -\nabla f(\boldsymbol{x}_k). \tag{1}$$

In our case we have

$$\nabla f = \left(3(x-1)^2, 2y\right),$$

and

$$\nabla^2 f = \begin{pmatrix} 6(x-1) & 0 \\ 0 & 2 \end{pmatrix}.$$

Solving the linear system (1) for $x_k \neq 1$ we obtain

$$\boldsymbol{p}_{k+1} = -\left(\frac{x_k - 1}{2}, y_k\right).$$

Newton's method with unit step yields

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{p}_{k+1} = \left(\frac{x_k + 1}{2}, 0\right).$$

We will now use induction to prove that

$$(x_k, y_k) = \left(1 + \frac{1}{2^k}(x_0 - 1), 0\right). \tag{2}$$

We first prove (2) for the case $k := 1$.

$$(x_1, y_1) = \left(\frac{x_0 + 1}{2}, 0\right) = \left(1 + \frac{x_0 - 1}{2}, 0\right).$$

We now assume that (2) holds for the case $k := n$, and prove that it then holds for $k := n + 1$.

$$\begin{aligned} (x_{n+1}, y_{n+1}) &= \left(\frac{x_n + 1}{2}, 0\right) \\ &= \left(\frac{1}{2} + \frac{1}{2}\left(1 + \frac{1}{2^n}(x_0 - 1)\right), 0\right) \\ &= \left(1 + \frac{1}{2^{n+1}}(x_0 - 1), 0\right). \end{aligned}$$

Thus we have shown that (2) holds for all $k \in \mathbb{N}$.

Newton's method converges to the point $(1, 0)$ which is neither a global nor a local optimum since $-\epsilon^3 = f(1 - \epsilon, 0) < f(1, 0) = 0$ for all $\epsilon > 0$. This contradicts the definition of local optimality.

Question 3

(sufficient global optimality conditions)

(1p) a) This is Theorem 4.3.

(2p) b) This is Theorem 5.45.

Question 4

(convexity)

(1p) a) The claim is true. Clearly, x_1^4 is a convex function, and since we know that a sum of convex functions remains convex, what is left to check is if $x_2^2 + 4x_2x_3 + 5x_3^2 := g(x_2, x_3)$ is convex. A computation of the eigenvalues to the hessian of g shows that they are $\lambda = 6 \pm \sqrt{32} > 0$. Therefore, the hessian is positive semidefinite for all $\boldsymbol{x} \in \mathbb{R}^3$ and thus, g is convex. We conclude that f is convex.

(1p) b) The claim is true. Let $h(\boldsymbol{x}) := 2x_1 - x_2$ and $g(\boldsymbol{x}) := x_2^2$ and observe that they are both convex. f is not differentiable so we cannot use the same procedure as in a); instead we use the definition. Let $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ be two arbitrary points and let $\lambda \in (0,1)$. Since h and g are convex, we have that

$$h(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leq \lambda h(\boldsymbol{x}^1) + (1-\lambda)h(\boldsymbol{x}^2) \text{ and}$$
$$g(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leq \lambda g(\boldsymbol{x}^1) + (1-\lambda)g(\boldsymbol{x}^2).$$

Therefore,

$$f(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) = \max\left\{h(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2), g(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2)\right\} \leq$$
$$\max\left\{\lambda h(\boldsymbol{x}^1) + (1-\lambda)h(\boldsymbol{x}^2), \lambda g(\boldsymbol{x}^1) + (1-\lambda)g(\boldsymbol{x}^2)\right\} \leq$$
$$\max\left\{\lambda h(\boldsymbol{x}^1), \lambda g(\boldsymbol{x}^1)\right\} + \max\left\{(1-\lambda)h(\boldsymbol{x}^2), (1-\lambda)g(\boldsymbol{x}^2)\right\} =$$
$$\lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2),$$

where the last inequality comes from the obvious fact that

$$\max\{a+b, c+d\} \leq \max\{a,c\} + \max\{b,d\}.$$

(1p) c) The claim is false. The hessian to f is given by

$$\nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 12x_1 + 2x_2^2 & 4x_1x_2 \\ 4x_1x_2 & 12x_2^2 + 8 \end{pmatrix}$$

and we conclude that its eigenvalues at $\boldsymbol{x} = (0,0)^{\mathrm{T}}$ are $\lambda_1 = 8$ and $\lambda_2 = 0$, i.e., the matrix is positive semidefinite but not positive definite. Therefore we cannot conclude anything about the local convexity from this fact. But now look at the line given by

$$\begin{cases} x_1 = t \\ x_2 = 0 \end{cases} \quad \text{and let } \boldsymbol{x}^1 = \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}, \ \boldsymbol{x}^2 = \begin{pmatrix} -\varepsilon \\ 0 \end{pmatrix}.$$

We have $f(\boldsymbol{x}^1) = 2\varepsilon^3$, $f(\boldsymbol{x}^2) = -2\varepsilon^3$ and $f(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) = 2\varepsilon^3(2\lambda-1)^3$. Therefore, we get that $f(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) > \lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2)$ when $(2\lambda - 1)^3 > 2\lambda - 1$ which is true for all $\lambda < 1/2$. This counter-example shows that f is not locally convex around the origin.

(3p) Question 5

(linear programs)

The set in question is described by primal–dual feasibility, and the inverse of weak duality. As the first two parts of the constraints describe weak duality, in total the system describes strong duality:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &\geq \boldsymbol{b}, \\ \boldsymbol{x} &\geq \boldsymbol{0}^n, \\ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} &\leq \boldsymbol{c}, \\ \boldsymbol{y} &\geq \boldsymbol{0}^m, \\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} &\leq \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \end{aligned}$$

(3p) Question 6

modelling

Let x_{ij} be the flow sent through edge $(i,j) \in E$. Introduce an auxiliary variable y_{ij} for each edge $(i,j) \in E$, which is the cost associated with edge $(i,j) \in E$.

Then the problem can be formulated as

$$
\begin{aligned}
\text{minimize} \quad & \sum_{(i,j)\in E} y_{ij}, \\
\text{subject to} \quad & y_{ij} \geq m_{ij}^l + k_{ij}^l x_{ij}, && (i,j) \in E,\ l = 1, \dots, n, \\
& \sum_{j:(i,j)\in E} x_{ij} = \sum_{j:(j,i)\in E} x_{ji}, && i \in V \setminus \{s,t\}, \\
& \sum_{j:(s,j)\in E} x_{sj} = d, \\
& \sum_{j:(j,t)\in E} x_{jt} = d, \\
& 0 \leq x_{ij} \leq c_{ij}, && (i,j) \in E.
\end{aligned}
$$

(3p) Question 7

Lagrangian duality

a) Lagrangian relaxing the constraint yields the Lagrangian function $\mathcal{L}(\boldsymbol{x}, \mu) = x_1^2 - 4x_1 + 2x_2^2 + \mu(2x_1 + x_2 - 2)$. From the stationary conditions for the Lagrangian, we get that

$$x_1(\mu) = \begin{cases} 2 - \mu, & 0 \leq \mu \leq 2, \\ 0, & \mu \geq 2; \end{cases} \qquad x_2(\mu) = 0,\ \mu \geq 0.$$

We then get the following expression for the Lagrangian dual function, to be maximized over $\mu \geq 0$.

$$q(\mu) = \begin{cases} -\mu^2 + 2\mu - 4, & 0 \leq \mu \leq 2, \\ -2\mu, & \mu \geq 2, \end{cases}$$

Calculating the derivative of q, we get

$$q'(\mu) = \begin{cases} -2\mu + 2, & 0 \leq \mu \leq 2, \\ -2, & \mu \geq 2, \end{cases}$$

It is clear that q is concave and differential for every $\mu \geq 0$. It is in fact strictly concave.

b) Setting $q'(\mu) = 0$ as a first guess, we obtain $q'(\mu) = 0$ for $\mu = 1$. Since the dual problem is convex this is the optimal solution, i.e., $\mu^* = 1$, with objective value $q(\mu^*) = -3$.

c) The Lagrangian optimal solution in $\boldsymbol{x}$ for $\mu = \mu^*$ is, from a), $\boldsymbol{x} = (1, 0)^{\mathrm{T}}$. $\boldsymbol{x}$ is feasible and $f(\boldsymbol{x}) = q(\mu^*)$, so by weak duality, $\boldsymbol{x}$ is an optimal solution. According to duality theory for convex problems over polyhedral sets, all primal optimal solutions are generated from Lagrangian optimal solutions given an optimal dual vector. Since $\boldsymbol{x}(\mu^*)$ here is the unique vector $\boldsymbol{x} = (1, 0)^{\mathrm{T}}$ this must also be the unique optimal solution to the primal problem.