Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 18–01–09
Examiner: Michael Patriksson

Question 1

(the simplex method)

(1p) a) Rewrite the problem into standard form by letting $x_1 := x_1^+ - x_1^-$ and adding/subtracting slack variables s_1 and s_2 to the left-hand side in the first and second constraint, respectively. Moreover, let $z := -z$ to get the problem on minimization form. Thus, we get the following linear program:

$$\begin{array}{llrrrrrl} \text{minimize} & z = & x_1^+ & - x_1^- & + 2x_2, & & & \\ \text{subject to} & & -x_1^+ & + x_1^- & + x_2 & + s_1 & & = 5, \\ & & & & x_2 & & - s_2 & = 2, \\ & & x_1^+, & x_1^-, & x_2, & s_1, & s_2 & \geq 0. \end{array}$$

(2p) b) Introducing the artificial variable a, phase I gives the problem

$$\begin{array}{llrrrrrrl} \text{minimize} & w = & a, & & & & & & \\ \text{subject to} & & -x_1^+ & + x_1^- & + x_2 & + s_1 & & & = 5, \\ & & & & x_2 & & - s_2 & + a & = 2, \\ & & x_1^+, & x_1^-, & x_2, & s_1, & s_2, & a & \geq 0. \end{array}$$

Using the starting basis $(s_1, a)^T$ gives

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The reduced costs, $\bar{\boldsymbol{c}}_N^T = \boldsymbol{c}_N^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{N}$, for this basis is $\bar{\boldsymbol{c}}_N^T = \begin{pmatrix} 0 & 0 & -1 & 1 \end{pmatrix}$, which means that x_2 enters the basis. The minimum ratio test implies that a leaves.

Thus, we move on to phase II using the basis $(s_1, x_2)^T$, and

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

The new reduced costs are $\bar{\boldsymbol{c}}_N^T = \begin{pmatrix} 1 & -1 & 2 \end{pmatrix}$ which means that x_1^- enters the basis. The minimum ratio test implies that s_1 leaves.

Updating the basis, now with $(x_1^-, x_2)^T$, gives

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

The new reduced costs are $\bar{\boldsymbol{c}}_N^T = \begin{pmatrix} 0 & 1 & 3 \end{pmatrix}$ which means that the current basis is optimal. The optimal solution is thus
$\boldsymbol{x}^* = \begin{pmatrix} x_1^+ & x_1^- & x_2 & s_1 & s_2 \end{pmatrix}^T = \begin{pmatrix} 0 & 3 & 2 & 0 & 0 \end{pmatrix}^T$
with optimal objective function value $z^* = 1$.

Question 2

(Lagrangian duality and convexity)

(2p) a) We create the Lagrangian function

$$L(\boldsymbol{x}, \mu) = (x_1-1)^2-2x_2+\mu(2x_2-x_1-2) = \left(x_1^2 - 2x_1 - \mu x_1\right)+2(\mu-1)x_2+1-2\mu. \tag{1}$$

The dual function then is

$$q(\mu) = \min_{\boldsymbol{x}\geq 0} L(x, \mu) = 1-2\mu + \min_{x_1 \geq 0} \left(x_1^2 - 2x_1 - \mu x_1\right) + \min_{x_2\geq 0} 2(\mu-1)x_2. \tag{2}$$

At $\mu = 0$, since the objective function coefficient for x_2 is negative, letting $x_2 \to \infty$ yields unbounded solutions to the Lagrangian subproblem. Thus $q(0) = -\infty$. At $\mu = 2$, to minimize the convex quadratic problem in x_1 we let $x_1 = 1 + \mu/2 = 2$, and $x_2 = 0$. Thus $q(2) = -7$. By weak duality it follows that $q(2) \leq f^*$. To find an upper bound, choose any feasible point, e.g. $(x_1, x_2) = (1, 1)$, which has objective value -2. Hence $f^* \in [-7, -2]$.

(1p) b) See course book.

Question 3

(Karush-Kuhn-Tucker)

(1p) a) Let $g_1(\boldsymbol{x}) := x_1 + x_2 - 5$, $g_2(\boldsymbol{x}) := -x_1$ and $g_3(\boldsymbol{x}) := -x_2$, with respective gradients $(1,1)^T$, $(-1,0)^T$ and $(0,-1)^T$.
Moreover, $\nabla f = (-2(x_1-3), -2(x_2-1))^T$. The KKT-conditions are as follows:

$$-2(x_1-3) + \mu_1 - \mu_2 = 0,$$
$$-2(x_2-1) + \mu_1 - \mu_3 = 0,$$
$$\mu_1, \mu_2, \mu_3 \geq 0,$$
$$x_1 + x_2 - 5 \leq 0,$$
$$-x_1 \leq 0,$$
$$-x_2 \leq 0,$$
$$\mu_1(x_1 + x_2 - 5) = 0,$$
$$\mu_2(-x_1) = 0,$$
$$\mu_3(-x_2) = 0.$$

Since the functions g_i, $i = 1, 2, 3$, are convex and there exists an inner point (for example $(1,1)^T$), the problem satisfies Slater CQ. Thus, the KKT-conditions are necessary.

(2p) b) By visually analyzing the figure, we can see that there is a total of 7 KKT-points. To find all of them analytically, let different combinations of constraints be active and solve for $\boldsymbol{x}$ in the KKT-conditions.

For instance, let g_1 be the only active constraint. Then, $x_1 + x_2 - 5 = 0$ and $\mu_2 = \mu_3 = 0$. This, together with the first two KKT-conditions, gives that $x_1 = \frac{7}{2}$ and $x_2 = \frac{3}{2}$. Thus, we get the KKT-point $\boldsymbol{x}^1 = (\frac{7}{2}, \frac{3}{2})^T$.

Similar calculations for other active constraints gives the KKT-points $\boldsymbol{x}^2 = (3,0)^T$, $\boldsymbol{x}^3 = (0,1)^T$, $\boldsymbol{x}^4 = (5,0)^T$, $\boldsymbol{x}^5 = (0,5)^T$, $\boldsymbol{x}^6 = (0,0)^T$ and $\boldsymbol{x}^7 = (3,1)^T$. Note that $\boldsymbol{x}^7$ is found when there are no active constraints, i.e. an inner point where $\nabla f(\boldsymbol{x}) = 0$.

Since the KKT-conditions are necessary, the global optimum must be in at least one KKT-point. Trying all of them gives $f^* = -25$ for $\boldsymbol{x}^* = \boldsymbol{x}^5 = (0,5)^T$.

Question 4

(unconstrained optimization)

We have that

$$\nabla f(\boldsymbol{x}) = (2x_1 + 2x_2 + 4, 2x_1 - 4x_2)^{\mathrm{T}}, \quad \nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 2 & 2 \\ 2 & -4 \end{pmatrix} \tag{1}$$

a) For the steepest descent method:

$$\boldsymbol{p} = -\nabla f(\bar{\boldsymbol{x}}) = (-4, 0)^{\mathrm{T}} \tag{2}$$

b) For Newtons method:

$$\boldsymbol{p} = -\left[\nabla^2 f(\bar{\boldsymbol{x}})\right]^{-1} \nabla f(\bar{\boldsymbol{x}}) = (-4/3, -2/3)^{\mathrm{T}} \tag{3}$$

c) For Newtons method with Levenberg-Marquardt modification:

$$\boldsymbol{p} = -\left[\nabla^2 f(\bar{\boldsymbol{x}}) + \gamma I\right]^{-1} \nabla f(\bar{\boldsymbol{x}}) = (-4/9, 2/9)^{\mathrm{T}} \tag{4}$$

The methods a) and c) always finds descent directions (if γ is chosen large enough).

(3p) Question 5

(modelling)

A suggested integer programming formulation is as follows:

Sets:
$\mathcal{L} := \{i | i \in \{1, ..., 7\}\}$, the set of wind turbines,
$\mathcal{M} := \{j | j \in \{Mon, ..., Fri\}\}$, the set of different days,
$\mathcal{N} := \{k | k \in \{1, 2\}\}$, the set of two maintenance teams.

To simplify the problem, we add a parameter c_{ij} $i \in \mathcal{L}$, $j \in \mathcal{M}$, are the maintenance cost for different wind turbines at each day.

The decision variables are:

$$x_{i,j,k} = \begin{cases} 1 & \text{if maintenance team } k \in \mathcal{N} \text{ maintain wind turbine } i \in \mathcal{L} \text{ at day } j \in \mathcal{M}, \\ 0 & \text{otherwise.} \end{cases}$$

Model:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in \mathcal{L}} \sum_{j \in \mathcal{M}} \sum_{k \in \mathcal{N}} c_{ji} x_{ijk}, \\ \text{subject to} \quad & \sum_{j \in \mathcal{M}} \sum_{k \in \mathcal{N}} x_{ijk} = 1 && i \in \mathcal{L}, \\ & \sum_{i \in \mathcal{L}} x_{ijk} \leq 1 && k \in \mathcal{N}, j \in \mathcal{M}, \\ & x_{ijk} \in \{0, 1\} && i \in \mathcal{L}, j \in \mathcal{M}, k \in \mathcal{N}. \end{aligned}$$

Question 6

(true or false)

(1p) a) The claim is false. The functions h_i, $i = 1, \dots, k$ defining the equality constraints must be affine.

(1p) b) The claim is true.
Choose arbitrary two points, $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$, an $\alpha \in [0,\ 1]$,

$$\begin{aligned} & \alpha f(\boldsymbol{x}^1) + (1-\alpha) f(\boldsymbol{x}^2) \\ = & \alpha \ln \sum_{j=1}^n e^{a_j x_j^1} + (1-\alpha) \ln \sum_{j=1}^n e^{a_j x_j^2} \\ = & \ln \sum_{j=1}^n e^{a_j x_j^1 \alpha} + \ln \sum_{j=1}^n e^{a_j x_j^2 (1-\alpha)} \\ = & \ln \sum_{j=1}^n e^{a_j x_j^1 \alpha} \sum_{j=1}^n e^{a_j x_j^2 (1-\alpha)} \qquad \text{since } e^x > 0, \forall x \in \mathbb{R} \\ \geq & \ln \sum_{j=1}^n e^{a_j x_j^1 \alpha} e^{a_j x_j^2 (1-\alpha)} \\ = & \ln \sum_{j=1}^n e^{a_j (x_j^1 \alpha + x_j^2 (1-\alpha))} \\ = & f(\alpha \boldsymbol{x}^1 + (1-\alpha) \boldsymbol{x}^2) \end{aligned}$$

By definition, f is a convex function.

(1p) c) The claim is false. Consider the linear program to minimize x_2 subject to the constraints $0 \leq x_j \leq 4, j = 1, 2$, and the additional constraint that $x_1 + x_2 \leq 2$. This problem has the optimal solution set $X^* = \{x \in \mathbb{R}^2 | x_1 \in [0, 2]; x_2 = 0\}$. At the optimal solution $x^* = (1, 0)^{\mathrm{T}}$, $x_1 + x_2 < 2$ holds. Believing that this means that the constraint $x_1 + x_2 \leq 2$ therefore is redundant results, however, in a grave mistake, as the new problem has the optimal set $X^*_{\text{new}} = \{x \in \mathbb{R}^2 | x_1 \in [0, 4]; x_2 = 0\}$.

Question 7

(LP duality)

See Theorem 10.6 in the course book.
