Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 17–08–24
Examiner: Michael Patriksson

Question 1

(simplex method)

(2p) a) Rewrite the problem into standard form by adding slack variables s_1 and s_1 to the left-hand side in the first and second constraint, respectively. Thus, we get the following linear program:

$$\begin{array}{lrcrcrcrcrcl}
\text{maximize} & z = 4x_1 & + & 2x_2 & + & 2x_3, & & & & & & \\
\text{subject to} & x_1 & - & x_2 & + & 2x_3 & + & s_1 & & & = & 2, \\
 & 2x_1 & + & x_2 & + & x_3 & & & + & s_2 & = & 8, \\
 & x_1, & & x_2, & & x_3, & & s_1, & & s_2 & \geq & 0.
\end{array}$$

We start directly with phase II at the origin, using the starting basis $(s_1, s_2)^T$. This iteration,

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 2 \\ 8 \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}.$$

The reduced costs, $\bar{\boldsymbol{c}}^T = \boldsymbol{c}_N^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{N}$, for this basis is $\bar{\boldsymbol{c}}^T = \begin{pmatrix} 4 & 2 & 2 \end{pmatrix}$, which means that x_1 enters the basis. The minimum ratio test implies that s_1 leaves.

Updating the basis, we now have $(x_1, s_2)^T$ in the basis and

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

The new reduced costs are $\bar{\boldsymbol{c}}^T = \begin{pmatrix} -4 & 6 & -6 \end{pmatrix}$ which means that x_2 enters the basis. The minimum ratio test implies that s_2 leaves.

Once again updating the basis, now with $(x_1, x_2)^T$, gives

$$\boldsymbol{B} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 3\frac{1}{3} \\ 1\frac{1}{3} \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

The new reduced costs are $\bar{\boldsymbol{c}}^T = \begin{pmatrix} 0 & -2 & 0 \end{pmatrix}$ which means that the current basis is optimal. The optimal solution is thus
$\boldsymbol{x}^* = \begin{pmatrix} x_1 & x_2 & x_3 & s_1 & s_2 \end{pmatrix}^T = \begin{pmatrix} 3\frac{1}{3} & 1\frac{1}{3} & 0 & 0 & 0 \end{pmatrix}^T$
with optimal objective function value $z^* = 16$.

(1p) b) The LP dual is

$$\begin{array}{lrcrcl}
\text{minimize} & w = 2y_1 & + & 8y_2, & & \\
\text{subject to} & y_1 & + & 2\,y_2 & \geq & 4, \\
& -y_1 & + & y_2 & \geq & 2, \\
& 2y_1 & + & y_2 & \geq & 2, \\
& y_1, & & y_2 & \geq & 0.
\end{array}$$

Drawing this problem, it easy to see that the optimal solution is $\boldsymbol{y}^* = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ with optimal objective function value $w^* = 16$.

(Note: Since there exists an optimal solution to the primal problem, strong duality actually implies that the dual problem also has an optimal solution.)

(3p)
Question 2

(Lagrangian duality)

Introducing the scalar $\lambda \in \mathbb{R}$, the KKT conditions state that

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \qquad \text{and} \qquad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1.$$

The interpretation is that the (primal) solutions in $\boldsymbol{x}$ are eigenvectors of the matrix $\boldsymbol{A}$, and that the (dual) variable λ is an eigenvalue.

See also Example 5.54 from the course book.

(3p) Question 3

(Newton's method)

(1p) a) Newton's equation:

$$x_{k+1} = x_k - \frac{\alpha {x_k}^{\alpha-1} - e^{x_k}}{\alpha(\alpha-1){x_k}^{\alpha-2} - e^{x_k}}$$

(1p) b) For example, take $x_0 = 1$, $\alpha = 4$. Then we can get $x_1 = 0.8619$, $x_2 = 0.8323$, $x_3 = 0.8310$, $x_4 = 0.8310$, ... These initial values cause the Newtons method to generate a sequence stuck at $x = 0.8310$ which is the local minimum.

(1p) c) The objective function of the problem is not convex in general [may be verified by analyzing the sign of the Hessian $\alpha(\alpha-1)x^{\alpha-2} - e^x$]. Since the convergence of the Newton method is local in nature, the method is most likely to converge to the nearest local minimum (or maximum if the Hessian is negative definite). The engineer thus wrongly assumes the global convergence of the Newton method on non convex functions.

Question 4

(unconstrained optimization)

The main problem in this question lies in the fact that we need to cope with the fact that the value of $\nabla f(\boldsymbol{x})$ needs to be exactly zero in order to conclude that $\boldsymbol{x}$ is stationary. Hence the exact measure

$$\nabla f(\boldsymbol{x}) = \mathbf{0}^n$$

needs to be replaced by a sensible tolerance. The course book include, in Section 11.5, a list of three combinations of criteria, based on a small norm of the gradient of f, a small decrease in the value of f between two iterates in relation to the size of problem data, and a small shift in the vector $\boldsymbol{x}$ between iterations.

(3p) Question 5

(modelling)

The decision variables are:

$$x_{i,j} = \begin{cases} 1 & \text{if person } i \in G_1 \text{ and person } j \in G_2 \text{ pair up,} \\ 0 & \text{otherwise.} \end{cases}$$

Model:

$$\begin{aligned} \text{maximize} \quad & \sum_{i \in G_1} \sum_{j \in G_2} (a_{ij} + b_{ji}) x_{ij}, \\ \text{subject to} \quad & \sum_{i \in G_1} x_{ij} = 1 && j \in G_2, \\ & \sum_{j \in G_2} x_{ij} = 1 && i \in G_1, \\ & (a_{ij} - p) x_{ij} \geq 0 && i \in G_1, j \in G_2, \\ & (b_{ji} - p) x_{ij} \geq 0 && i \in G_1, j \in G_2, \\ & x_{i,j} \in \{0, 1\} && i \in G_1, j \in G_2, . \end{aligned}$$

Question 6

(true or false)

(1p) a) The claim is false in general. (However, if f is lower bounded, then it has a minimum.)

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

The quadratic form $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} - \boldsymbol{b}^T \boldsymbol{x}$ has no minimum.

(1p) b) The claim is false, as a linear program may have no feasible solutions.

(1p) c) The claim is true, and is established in Theorem 6.4 in the textbook.

Question 7

(a simple optimization problem)

The KKT conditions for this problem amount, apart from complementary and primal feasibility, to finding a solution in the pair $(\boldsymbol{x}, \mu)^T \in R^n \times R_+$ to the nonlinear equations formed by the stationarity conditions for the Lagrangian with respect to $\boldsymbol{x}$, that is, for all $j = 1, ..., n$,

$$\frac{a_j}{x_j^2} + \frac{\mu}{x_j} = 0$$

This is clearly impossible, as $x_j > 0$ must be fulfilled, and $a_j > 0$ holds. We therefore conclude that there are no KKT points for this problem.

Can there be optimal solutions that are not KKT points? No, because the linear independence CQ (LICQ) is fulfilled for this problem, so the KKT conditions are necessary conditions for both local and global optimal solutions.