Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date: 16–04–05
Examiner: Michael Patriksson

Question 1

(the simplex method)

(1p) a) The modified problem is always feasible by construction. For example, a feasible solution is $x_i = 0$ for $i = 1, 2, 3, 4$ and $y_1 = 5$ and $y_2 = 3$. Assuming that the modified problem has optimal objective value bounded from below, the modified problem always has finite optimal solution. Let $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ denote the x-part and y-part of the optimal solution, respectively. Depending on the value of $\boldsymbol{y}^*$, two cases are possible:

- At optimality, $y_1^* = y_2^* = 0$. In this case, the original problem is feasible. In addition, $\boldsymbol{x}^*$ is an optimal solution to the original problem. It is obvious that $\boldsymbol{x}^*$ is feasible to the original problem. If there were some $\tilde{x}$ feasible to the original problem with an objective value smaller than that of $\boldsymbol{x}^*$, then $\tilde{\boldsymbol{x}}$ together with $\boldsymbol{y}^* = \boldsymbol{0}$ form a better feasible solution to the modified problem. This contradicts the optimality of x^* and $\boldsymbol{y}^*$ for the modified problem.
- At optimality, at least one of y_1^* and y_2^* is positive. In this case, the original problem is infeasible. If a vector $\tilde{\boldsymbol{x}}$ were feasible to the original problem, then $\tilde{\boldsymbol{x}}$ together with $\boldsymbol{y} = \boldsymbol{0}$ result in a better feasible solution of the modified problem than $\boldsymbol{x}^*$ with $\boldsymbol{y}^*$ (cf. the property of M). This would contradicts the optimality of $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ for the modified problem.

(2p) b) We can start the simplex method with y_1 and y_2 being the basic variables. The non-basic variables are x_1, x_2, x_3 and x_4.

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{c}_N^{\mathrm{T}} = \begin{pmatrix} 8 & 3 & 4 & 1 \end{pmatrix}, \quad \boldsymbol{c}_B^{\mathrm{T}} = \begin{pmatrix} M & M \end{pmatrix}$$
$$N = \begin{pmatrix} 2 & 1 & 3 & -1 \\ 1 & 1 & 2 & -1 \end{pmatrix}, \quad x_B = B^{-1}b = \begin{pmatrix} 5 \\ 3 \end{pmatrix}.$$

The reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N = \begin{pmatrix} 8 - 3M & 3 - 2M & 4 - 5M & 1 + 2M \end{pmatrix}.$$

We choose the third non-basic variable (i.e., x_3) to enter the basis, because it has the most negative reduced cost. The corresponding search direction for the basic variables are $d_B = -B^{-1}N_3 = (-3, -2)^{\mathrm{T}}$. The minimum ratio test indicates that

$$2 = \operatorname{argmin}\{\frac{5}{3}, \frac{3}{2}\},$$

and hence the second basic variable (i.e., y_2) leaves the basis.

At iteration two, we have x_3 and y_1 being the basic variables. The non-basic variables are x_1, x_2, x_4 and y_2.

$$B = \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{3}{2} \end{pmatrix}, \quad \boldsymbol{c}_N^{\mathrm{T}} = \begin{pmatrix} 8 & 3 & 1 & M \end{pmatrix}, \quad \boldsymbol{c}_B^{\mathrm{T}} = \begin{pmatrix} 4 & M \end{pmatrix}$$
$$N = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 1 & 1 & -1 & 1 \end{pmatrix}.$$

The reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N = \begin{pmatrix} 6 - \frac{M}{2} & 1 + \frac{M}{2} & 3 - \frac{M}{2} & -2 + \frac{3M}{2} \end{pmatrix}.$$

We choose the third non-basic variable (i.e., x_4) to enter the basis. The corresponding search direction for the basic variables are $d_B = -B^{-1}N_3 = (\frac{1}{2}, -\frac{1}{2})^{\mathrm{T}}$. Therefore, the second basic variable (i.e., y_1) leaves the basis.

At iteration three, we have basic variables being x_3 and x_4. The non-basic variables are x_1, x_2, y_1 and y_2.

$$B = \begin{pmatrix} 3 & -1 \\ 2 & -1 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & -1 \\ 2 & -3 \end{pmatrix}, \quad \boldsymbol{c}_N^{\mathrm{T}} = \begin{pmatrix} 8 & 3 & M & M \end{pmatrix}, \quad \boldsymbol{c}_B^{\mathrm{T}} = \begin{pmatrix} 4 & 1 \end{pmatrix},$$
$$N = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \quad x_B = B^{-1}b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

The reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N = \begin{pmatrix} 3 & 4 & M - 6 & M + 7 \end{pmatrix}.$$

The reduced costs are all nonnegative. The simplex method terminates with optimal solution

$$x^* = (0, 0, 2, 1)^{\mathrm{T}}, \quad \boldsymbol{y}^* = (0, 0), \quad z^* = 9$$

As explained in part a), x^* is also an optimal solution to the original problem with objective value 9.

Question 2

(true or false)

(1p) a) Impossible to say, since the original problem may lack optimal solutions.

(1p) b) True—see Exercise 11.1.

(1p) c) Impossible to say, since the function f may not be convex.

(3p) Question 3

(optimality conditions)

This is Theorem 10.10.

(3p) Question 4

(Frank–Wolfe)

We can only guarantee that the point obtained is stationary. If f however is concave, then we establish that the point obtained is optimal.

(3p) Question 5

(Lagrangian duality)

This is Theorem 6.8.

(3p) Question 6

(integer programming modeling)

A suggested integer programming formulation is as follows: each square is labeled with an i nteger index (e.g., $1, \ldots, n^2$). For each square i, we define the neighborhood N_i to be the set of all indices of squares that can be attacked if a queen is placed at square i. For each i, we define a 0-1 binary decision variable $x_i \in \{0, 1\}$ such that a queen is placed at square i if and only if $x_i = 1$. Then, an integer program modeling t he desired queen configuration problem is

$$\begin{aligned}
\underset{\boldsymbol{x}}{\text{minimize}} \quad & \sum_{i=1}^{n^2} x_i \\
\text{subject to} \quad & x_i + \sum_{j \in N_i} x_j \geq 1, \quad i = 1, \ldots, n^2 \\
& (n^2 - 1)x_i + \sum_{j \in N_i} x_j \leq n^2 - 1, \quad i = 1, \ldots, n^2 \\
& x_i \in \{0, 1\}, \quad i = 1, \ldots, n^2.
\end{aligned}$$

In the model above, the first constraint specifies that for each square i either there is a queen or the square can be attacked by a queen in the neighborhood N_i. The second con straint specifies that if a queen is placed at square i, then no queen can be placed at a ny square in the neighborhood N_i (we can replace $n^2 - 1$ by any constant larger than tha t). The two constraints model exactly the conditions required by the queen configuration pr oblem.

(3p) Question 7

(gradient projection algorithm)

At $\boldsymbol{x}^0 = (0, 0)^{\mathrm{T}}$, the objective gradient vector is $\nabla f(\boldsymbol{x}^0) = (x_1 - 2, x_2 - \frac{3}{2})^{\mathrm{T}} = (-2, -\frac{3}{2})^{\mathrm{T}}$. Hence, the search direction is $\boldsymbol{p}^0 = -\nabla f(\boldsymbol{x}^0) = (2, \frac{3}{2})^{\mathrm{T}}$. Because of the form of the feasible set X (i.e., box constraints), projection on X can be expressed analytically. The projection arc is of the form (for $0 \leq \alpha^0 \leq 1$):

$$\mathrm{Proj}_X[\boldsymbol{x}^0 + \alpha^0 \boldsymbol{p}^0] = \begin{pmatrix} \min\{1, 0 + 2\alpha^0\} \\ \min\{1, 0 + \frac{3}{2}\alpha^0\} \end{pmatrix}.$$

Hence, the objective function (to be minimized) for exact line search is

$$\begin{aligned}
f^0(\alpha^0) \quad &:= \quad \tfrac{1}{2}(\min\{1, 2\alpha^0\} - 2)^2 + \tfrac{1}{2}(\min\{1, \tfrac{3}{2}\alpha^0\} - \tfrac{3}{2})^2 \\
&= \quad \begin{cases} \frac{1}{2}\Big(4(\alpha^0 - 1)^2 + \frac{9}{4}(\alpha^0 - 1)^2\Big) & 0 \leq \alpha^0 \leq \frac{1}{2} \\ \frac{1}{2}\Big(1 + \frac{9}{4}(\alpha^0 - 1)^2\Big) & \frac{1}{2} \leq \alpha^0 \leq \frac{2}{3} \\ \frac{5}{8} & \frac{2}{3} \leq \alpha^0 \leq 1 \end{cases}.
\end{aligned}$$

Minimizing f^0 with $0 \leq \alpha^0 \leq 1$ yields the minimizing α^0 to be greater than or equal to $2/3$. Hence, the next iterate is

$$\boldsymbol{x}^1 = \mathrm{Proj}_X[\boldsymbol{x}^0 + \alpha^0 \boldsymbol{p}^0] = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

It is claimed that $\boldsymbol{x}^1$ is an optimal solution. First, note that the objective gradient a t $\boldsymbol{x}^1 = (1,1)^{\mathrm{T}}$ is $\nabla f(\boldsymbol{x}^1) = (x_1 - 2, x_2 - \frac{3}{2})^{\mathrm{T}} = (-1, -\frac{1}{2})^{\mathrm{T}}$. At $\boldsymbol{x}^1$ the active constraints are $x_1 \leq 1$ and $x_2 \leq 1$ with constraint function gradients being $(1,0)^{\mathrm{T}}$ and $(0,1)^{\mathrm{T}}$, respectively. As a result, $-\nabla f(\boldsymbol{x}^1)$ is in the cone of the active constraint gradients. This implies that $\boldsymbol{x}^1$ is a KKT point. In addition, the optimization problem is convex with affine constraints. Hence, the KKT point $\boldsymbol{x}^1$ is indeed an optimal solution.