Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 15–08–27

Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) We first rewrite the problem on standard form. We introduce slack variables s_1 and s_2. Consider the following linear program:

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

By introducing an artificial variable a, we get the Phase I problem to

$$\begin{array}{lrcrcrcrcrcl} \text{minimize} & w = & & & & & & & & a, & & \\ \text{subject to} & x_1 & - & 2x_2 & & & + & s_2 & & & = & 1, \\ & x_1 & + & x_2 & - & s_1 & & & + & a & = & 1, \\ & x_1, & & x_2, & & s_1, & & s_2, & & a & \geq & 0. \end{array}$$

The starting basis is $(s_2, a)^{\mathrm{T}}$. Calculating the reduced costs for the non-basic variables x_1, x_2, and s_1 we obtain $\tilde{\mathbf{c}}_N = (-1, -1, 1)^{\mathrm{T}}$, meaning that x_1 enters the basis. From the minimum ratio test, we get that a leaves the basis.

Updating the basis we now have $(s_2, x_1)^{\mathrm{T}}$ in the basis meaning that $w^* = 0$ and the basis found is corresponding to a basic feasible solution of the original problem in the standard form, i.e., the Phase II problem.

Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (-3, 2)^{\mathrm{T}}$. meaning that x_2 enters the basis. From the minimum ratio test we get that x_1 leaves the basis.

Updating the basis we now have $(s_2, x_2)^{\mathrm{T}}$ in the basis. Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (3, -1)^{\mathrm{T}}$, meaning that s_1 enters basis. From the minimum ratio test we get that $\mathbf{B}^{-1}\mathbf{N}_{s_1} = (-1, 0)^{\mathrm{T}} \leq \mathbf{0}$, meaning that the problem is unbounded.

(1p) b) The primal problem is unbounded, implying that $\mathbf{c}^{\mathrm{T}}\mathbf{x}^* = -\infty$. From weak duality we have that $\mathbf{b}^{\mathrm{T}}\mathbf{y} \leq \mathbf{c}^{\mathrm{T}}\mathbf{x}^*$ for all feasible $\mathbf{y}$, meaning that the dual problem is infeasible.

(3p) Question 2

(linear inequalities)

Consider the linear program to

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & -c^{\mathrm{T}} x, \\ \text{subject to} \quad & Ax \leq b, \end{aligned} \tag{1}$$

and its standard form equivalence

$$\begin{aligned} \underset{x^+, x^-, s}{\text{minimize}} \quad & -c^{\mathrm{T}} x^+ + c^{\mathrm{T}} x^-, \\ \text{subject to} \quad & Ax^+ - Ax^- + s = b, \\ & x^+ \geq 0, \; x^- \geq 0, \; s \geq 0. \end{aligned} \tag{2}$$

The dual of (2) is to

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & b^{\mathrm{T}} p, \\ \text{subject to} \quad & A^{\mathrm{T}} p \leq -c, \\ & -A^{\mathrm{T}} p \leq c, \\ & p \leq 0, \end{aligned} \tag{3}$$

and (3) is equivalent to

$$\begin{aligned} \underset{y}{\text{maximize}} \quad & -b^{\mathrm{T}} y, \\ \text{subject to} \quad & A^{\mathrm{T}} y = c, \\ & y \geq 0. \end{aligned} \tag{4}$$

If statement (a) holds, then the objective of (1) and (2) is bounded from below by $-d$. Hence, there exists an optimal solution to the dual of (2), which is (3). Consequently, by strong duality (cf. Theorem 10.6 in the text) the optimal objective values of (3) and (4) are equal to that of (2), which is bounded from below by $-d$. This implies that, for (4), there exists a vector $y \geq 0$ such that $A^{\mathrm{T}} y = c$ and $-b^{\mathrm{T}} y \geq -d$ (i.e., $b^{\mathrm{T}} y \leq d$). This statement is the same as (b).

Conversely, if (b) holds then (3) has at least one feasible solution with an objective value bounded from below by $-d$. Hence, by weak duality (cf. Theorem 10.4 in the text) every x feasible to (1) (i.e., $Ax \leq b$) must satisfy $-c^{\mathrm{T}} x \geq -d$. This implies statement (a).

(3p) Question 3

(the Frank–Wolfe algorithm)

Figure 1 shows the feasible set of the problem (i.e., the polyhedron with thick black boundary lines) and some contours of the objective function. Th e optimal solution is denoted by $x^\star$ (i.e., the red dot in the figure). $x^{(k)}$ for $k = 0, 1, 2$ denotes iterates visited by the Frank-Wolfe algorithm.

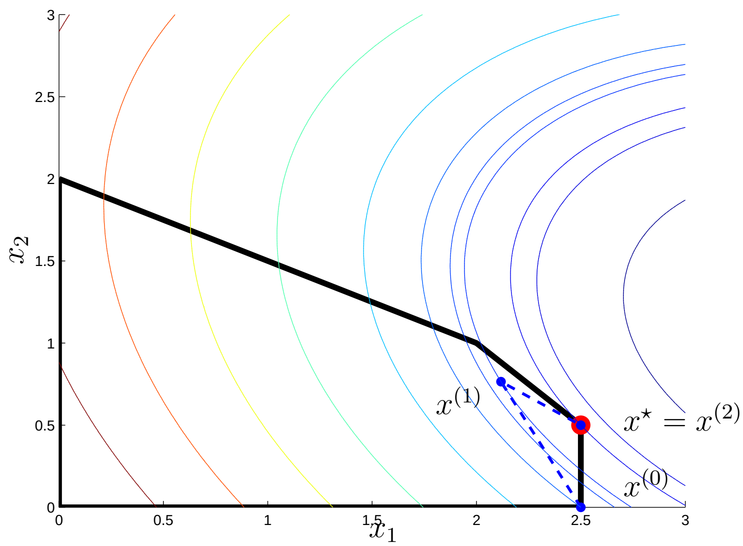

Figure 1: Illustration of the Frank-Wolfe algorithm. The feasible set is a polyhe dron with boundary denoted by the thick black lines. Some contours of the object ive function are shown. The optimal solution $x^\star = (2.5, 0.5)$. The dotted lines show the Frank-Wolfe iterations, with $x^{(k)}, k = 0, 1, 2$ denoting the i terates.

The details of the algorithm steps are as follows. Let X denote the feasible s et. Let $f(x_1, x_2)$ denote the objective function. For any given iterate $x^{(k)} = (x_1^{(k)}, x_2^{(k)})$. The objective function gradient vector is

$$\nabla f(x_1^{(k)}, x_2^{(k)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix}.$$

The search direction problem is

$$\underset{x \in X}{\text{minimize}} \quad \nabla f(x_1^{(k)}, x_2^{(k)})^{\mathrm{T}} x. \tag{1}$$

If $\min_{x \in X} \nabla f(x_1^{(k)}, x_2^{(k)})^{\mathrm{T}} x \geq \nabla f(x_1^{(k)}, x_2^{(k)})^{\mathrm{T}} x^{(k)}$, then by optimality conditions (for minimi zing a convex function over a convex feasible set) $x^{(k)}$ is optimal. Otherwi se, let $y^{(k)}$ denote an optimal solution to the search direction problem. Th en the exact minimization line search problem can be expressed into

$$\underset{\alpha \in [0,1]}{\text{minimize}} \quad f(\alpha x^{(k)} + (1-\alpha) y^{(k)}) \iff \underset{\alpha \in [0,1]}{\text{minimize}} \quad g\alpha^2 + h\alpha,$$

where

$$\begin{aligned} g &= \left(x^{(k)} - y^{(k)}\right)^{\mathrm{T}} \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} \left(x^{(k)} - y^{(k)}\right) \\ h &= \left(x^{(k)} - y^{(k)}\right)^{\mathrm{T}} \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} y^{(k)} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right). \end{aligned} \tag{2}$$

The minimizing value of α, denoted by $\alpha^{(k)}$, can be found using the optimality condition to be

$$\alpha^{(k)} = \begin{cases} 0 & \text{if} \quad -\frac{h}{2g} < 0 \\ -\frac{h}{2g} & \text{if} \quad 0 \leq -\frac{h}{2g} \leq 1 \\ 1 & \text{if} \quad -\frac{h}{2g} > 1 \end{cases}. \tag{3}$$

The iterate update formula is

$$x^{(k+1)} = \alpha^{(k)} x^{(k)} + (1 - \alpha^{(k)}) y^{(k)}. \tag{4}$$

Now we begin applying the Frank-Wolfe algorithm. At the first iteration with $x^{(0)} = (2.5, 0)$, the objective function gradient is

$$\nabla f(x_1^{(0)}, x_2^{(0)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \begin{bmatrix} -22 \\ -24 \end{bmatrix}.$$

To solve the search direction problem in (1), it is sufficient to restrict the feasible set to the set of all extreme points. That is,

$$\underset{x \in V}{\text{minimize}} \quad \nabla f(x_1^{(0)}, x_2^{(0)})^{\mathrm{T}} x, \tag{5}$$

where V is the set of all extreme points defined as

$$V = \Big\{(0,0), (0,2), (2,1), (2.5, 0.5), (2.5, 0)\Big\}.$$

This amounts to finding the minimum among five numbers: 0, -48, -68, -67, -55. The result is that $y^{(0)} = (2,1)$. Applying the formula in (**??**) yields

$$\begin{aligned} g &= \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)^{\mathrm{T}} \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right) &= 8.5 \\ h &= \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)^{\mathrm{T}} \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right) &= -4 \end{aligned}$$

According to (3), $\alpha^{(0)} = \frac{4}{17}$. Hence, by (4)

$$x^{(1)} = \frac{4}{17}(\frac{5}{2}, 0) + (1 - \frac{4}{17})(2, 1) = (\frac{36}{17}, \frac{13}{17}) \approx (2.12, 0.76).$$

This is shown in Figure 1.

At the next iteration with $x^{(1)} = (\frac{36}{17}, \frac{13}{17})$, we have

$$\nabla f(x_1^{(1)}, x_2^{(1)}) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \frac{1}{17} \begin{bmatrix} -400 \\ -200 \end{bmatrix} \approx \begin{bmatrix} -23.53 \\ -11.76 \end{bmatrix}.$$

Solving (5) amounts to finding the minimum of 0, −4, −10, −11, −10. This leads to $y^{(1)} = (2.5, 0.5)$. Applying (2) leads to

$$\begin{aligned} g &= \tfrac{1275}{1156} &\approx& \ 1.10 \\ h &= \tfrac{125}{34} &\approx& \ 3.68. \end{aligned}$$

Thus, according to (3) $\alpha^{(1)} = 0$, and from (4) $x^{(2)} = y^{(1)} = (2.5, 0.5)$ as shown in Figure 1.

At the final iteration with $x^{(2)} = (2.5, 0.5)$, we have

$$\nabla f(x_1^{(2)}, x_2^{(2)}) = \begin{bmatrix} -20 \\ -15 \end{bmatrix}.$$

Solving (5) leads to $y^{(2)} = x^{(2)} = (2.5, 0.5)$. Thus, it holds that

$$\min_{x \in X} \nabla f(x_1^{(2)}, x_2^{(2)})^{\mathrm{T}} x \geq \nabla f(x_1^{(2)}, x_2^{(2)})^{\mathrm{T}} x^{(2)}.$$

By optimality conditions, $x^{(2)} = (2.5, 0.5)$ is the optimal solution to our problem.

(3p) Question 4

(modelling)

The decision variables are:

$y_i = 1$, if school i is open, 0 otherwise

$x_{ij} = 1$, if students in area j attend school i, 0 otherwise

Model

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{10} c_i y_i + 2m \sum_{i=1}^{10} \sum_{j=1}^{J} b_j d_{ij} x_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{J} b_j x_{ij} \le k_i y_i, \quad i = 1, \dots, 10 \\
& \sum_{i=1}^{10} y_i \ge 7, \\
& \sum_{i=1}^{10} x_{ij} = 1, \qquad j = 1, \dots, J \\
& y_i \in \{0,1\}, \; i = 1, \dots, 10 \\
& x_{ij} \in \{0,1\}, \; i = 1, \dots, 10 \\
& \qquad j = 1, \dots, J
\end{aligned}
$$

The program is linear with integer variables.

Question 5

(true or false)

(1p) a) False. A simple example has $f(x) = x^2$ for $x \le 0$, and $x^3 + \mid x \mid$ for $x \ge 0$.

(1p) b) False. It provides a lower bound on the optimal value of the original (primal) problem.

(1p) c) True. Theorem 10.15 (necessary and sufficient conditions for global optimality) shows that an optimal dual solution is a vector of Lagrange multipliers.

Question 6

(interior penalty methods)

(1p) a) All functions involved are in C^1. The conditions on the penalty function are fulfilled, since $\phi'(s) = 1/s^2 \ge 0$ for all $s < 0$. Further, LICQ holds everywhere. The answer is yes.

(2p) b) With the given data, it is clear that the only constraint is (almost) fulfilled with equality: $(\boldsymbol{x}_6)_1^2 - (\boldsymbol{x}_6)_2 \approx -0.005422 \approx 0$. We set up the KKT conditions to see whether it is fulfilled approximately. Indeed, we have the following corresponding to the system $\nabla f(\boldsymbol{x}_6) + \hat{\mu}_6 \nabla g(\boldsymbol{x}_6) = \mathbf{0}^2$:

$$\begin{pmatrix} -6.4094265 \\ 3.39524 \end{pmatrix} + 3.385 \begin{pmatrix} 1.88778 \\ -1 \end{pmatrix} \approx \begin{pmatrix} -0.01929 \\ 0.01024 \end{pmatrix},$$

and the right-hand side can be considered near-zero. Since $\hat{\mu}_6 \geq 0$ we approximately fulfill the KKT conditions.

For the last part, we establish that the problem is convex. The feasible set clearly is convex, since g is a convex function and the constraint is on the "$\leq$"-form. The Hessian matrix of f is

$$\begin{pmatrix} 12(x_1 - 2)^2 + 2 & -4 \\ -4 & 8 \end{pmatrix},$$

which is positive semidefinite everywhere (in fact, positive definite outside of the region defined by $x_1 = 2$); hence, f is convex on $\mathbb{R}^2$. We conclude that our problem is convex, and hence the KKT conditions imply global optimality. The vector $\boldsymbol{x}_6$ therefore is an approximate global optimal solution to our problem.

Question 7

(the KKT conditions)

(2p) a) The KKT conditions are

$$\nabla f(\mathbf{x}) + \lambda \nabla h(\mathbf{x}) = \begin{pmatrix} x_2 + x_3 \\ x_1 + x_3 \\ x_1 + x_2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{0}.$$

There is only one feasible point fulfilling the KKT conditions:

$$\bar{\mathbf{x}} = (1, 1, 1)^{\mathrm{T}} \text{ with } \lambda = -2.$$

(1p) b) Since the eigenvalues of the Hessian of the objective function

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

are $\lambda_1 = \lambda_2 = -1, \lambda_3 = 2$ the objective function is not convex, indicating that the problem is unbounded. The KKT point $\bar{\mathbf{x}}$ is not an optimal solution.