Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 21–01–02
Examiner: Ann-Brith Strömberg

Note: the solutions presented here are brief in relation to the requirements on your answers, in particular regarding your motivations.

Question 1

(Simplex method)

(0.5p) a) The problem on standard form is:

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := -4x_1 + x_2, \\ \text{subject to} \quad & x_1 - x_2 + s_1 = 2, \\ & -x_1 + 2x_2 + s_2 = 1, \\ & x_1, \quad x_2, \quad s_1, \quad s_2 \geq 0. \end{aligned}$$

(1.5p) b) We can start directly in phase two since the slack variables provides an initial feasible basis.

First iteration: we have $x_B = (s_1, s_2), x_N = (x_1, x_2), B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $N = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}, c_B = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, c_N^{\mathrm{T}} = \begin{bmatrix} -4 & 1 \end{bmatrix}, B^{-1}b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Checking optimality:

$$\bar{c}_N^{\mathrm{T}} = c_N^{\mathrm{T}} - c_B^{\mathrm{T}} B^{-1} N = \begin{bmatrix} -4 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -4 & 1 \end{bmatrix}$$

Not optimal, minimum reduce costs indicate x_1 enter the basis.

Minimum ratio test: $B^{-1}N_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$\underset{i \in (B^{-1}N_1)_i > 0}{\operatorname{argmin}} \frac{(B^{-1}b)_i}{(B^{-1}N_1)_i} = \operatorname{argmin}\{\frac{2}{1}, -\}$$

hence, s_1 leaves the basis.

Second iteration: we have $x_B = (x_1, s_2), x_N = (x_2, s_1), B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, N = \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}, c_B = \begin{bmatrix} -4 \\ 0 \end{bmatrix}, c_N^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \end{bmatrix}, B^{-1}b = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Checking optimality:

$$\bar{c}_N^{\mathrm{T}} = c_N^{\mathrm{T}} - c_B^{\mathrm{T}} B^{-1} N = \begin{bmatrix} 1 & 0 \end{bmatrix} + \begin{bmatrix} 4 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 4 \end{bmatrix}$$

Not optimal, minimum reduce costs indicate x_2 enter the basis.

Minimum ratio test: $B^{-1}N_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$\underset{i \in (B^{-1}N_1)_i > 0}{\operatorname{argmin}} \frac{(B^{-1}b)_i}{(B^{-1}N_1)_i} = \operatorname{argmin}\{-, \frac{3}{1}\}$$

hence, s_2 leaves the basis.

Third iteration: we have $x_B = (x_1, x_2), x_N = (s_1, s_2), B = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}, B^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, N = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, c_B = \begin{bmatrix} -4 \\ 1 \end{bmatrix}, c_N^{\mathrm{T}} = \begin{bmatrix} 0 & 0 \end{bmatrix}, B^{-1}b = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$.

Checking optimality:

$$\bar{c}_N^{\mathrm{T}} = c_N^{\mathrm{T}} - c_B^{\mathrm{T}} B^{-1} N = \begin{bmatrix} 0 & 0 \end{bmatrix} + \begin{bmatrix} 7 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 3 \end{bmatrix} \geq 0$$

The solution in the original variables are $x_1 = 5, x_2 = 3$.

(1p) c) Continuing the third iteration, we have a new non-basic variable x_3.
$x_N = (x_3, s_1, s_2), N = \begin{bmatrix} 1 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}, c_N^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$.

Checking optimality:

$$\bar{c}_N^{\mathrm{T}} = c_N^{\mathrm{T}} - c_B^{\mathrm{T}} B^{-1} N = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 7 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 7 & 3 \end{bmatrix}$$

Not optimal, minimum reduce costs indicate x_3 enter the basis.

Minimum ratio test: $B^{-1}N_1 = \begin{bmatrix} -1 \\ -2 \end{bmatrix} \leq 0$, hence the problem is unbounded.

The ray of unboundedness in the original variables is $x_1 = 5 + t, x_2 = 3 + 2t, x_3 = t, t \geq 0$.

Question 2

(Farkas Lemma)

We have that there exists a vector $\boldsymbol{z} \leq \boldsymbol{0}$ such that $B\boldsymbol{z} - C\boldsymbol{z} = \boldsymbol{v}$. Which means that for $\boldsymbol{x} = -\boldsymbol{z}$ it holds that

$$(C - B)\boldsymbol{x} = \boldsymbol{v},$$
$$\boldsymbol{x} \geq \boldsymbol{0}.$$

Using Farkas lemma we then know that there can not exist any $\boldsymbol{u} \in \mathbb{R}^m$ such that

$$(C - B)^{\mathrm{T}}\boldsymbol{u} \geq \boldsymbol{0},$$
$$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u} < 0.$$

So there can not exist any $\boldsymbol{y} \in \mathbb{R}^m$ with $C^{\mathrm{T}}\boldsymbol{y} \leq B^{\mathrm{T}}\boldsymbol{y}$ and $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{y} > 0$.

(3p) Question 3

(KKT conditions)

(1p) a) Set $f(\boldsymbol{x}) = -c^t\boldsymbol{x}$, $g(\boldsymbol{x}) = \boldsymbol{x}^t\boldsymbol{x} - 1$. The KKT conditions are

$$\nabla f(\boldsymbol{x}) + \mu\nabla g(\boldsymbol{x}) = -c + 2\mu\boldsymbol{x},$$
$$\mu g(\boldsymbol{x}) = 0,$$
$$\mu \geq 0.$$

When $\bar{\boldsymbol{x}} = c/||c||$, $\mu = ||c||/2$, all the conditions are fulfilled. So $\bar{\boldsymbol{x}}$ is a KKT point.

(2p) b) Since the objective function and the feasible set are both convex, the problem is convex. Thus KKT conditions are sufficient. Since the feasible set is convex and $\mathbf{0}$ is an interior point, Slater CQ holds. Thus KKT conditions are necessary. To solve the KKT system, suppose $\tilde{\boldsymbol{x}}$ is a KKT point. If $g(\tilde{\boldsymbol{x}}) < 0$, then $\mu = 0$, but $\nabla f(\mathbf{x}) = c \neq \mathbf{0}$, contradiction. Thus $g(\tilde{\boldsymbol{x}}) = 0$, $\mu > 0$. $\tilde{\boldsymbol{x}} = c/2\mu$, plug it into $g(\tilde{\boldsymbol{x}}) = 0$, we get $\tilde{\boldsymbol{x}} = c/||c||$. So, $\bar{\boldsymbol{x}}$ is an unique KKT point. Since KKT conditions are both necessary and sufficient, $\bar{\boldsymbol{x}}$ is an unique global optimal.

(3p) Question 4

(Gradient projection)

Iteration 1: We have $\nabla f(\boldsymbol{x}^0) = (-2, -3)^T$. We need to project the point $(0, 0)^T - (-2, -3)^T = (2, 3)^T$ on the feasible region X. We graphically see that this projection is obtained by taking the point $(2, 2)$. Hence, $\boldsymbol{x}^1 = (2, 2)^T$.

Iteration 2: We have $\nabla f(\boldsymbol{x}^1) = (-2, 1)^T$. We need to project the point $(2, 2)^T - (-2, 1)^T = (4, 1)^T$ on the feasible region X. We graphically see that this projection is obtained by taking the point $(3, 1)$. Hence, $\boldsymbol{x}^2 = (3, 1)^T$.

The obtained point is neither a global nor a local minimum. This can be checked by, e.g., the KKT conditions and realizing that the point is not a stationary point.

(3p) Question 5

(modeling)

(1.5p) a) Definitions of additional sets

- $I := \{1, \ldots, 9\}$ be the index set of rows.
- $J := \{1, \ldots, 9\}$ be the index set of columns.
- $L := \{1, \ldots, 9\}$ be the index set of cells.
- $K := \{1, \ldots, 9\}$ be the index set of numbers.

The set of feasible solution S to the Sudoku is defined by:

$$\begin{aligned}
\sum_{i \in I} x_{ijk} &= 1, & j \in J, k \in K, \\
\sum_{j \in J} x_{ijk} &= 1, & i \in I, k \in K, \\
\sum_{(i,j) \in C_l} x_{ijk} &= 1, & l \in L, k \in K, \\
\sum_{k \in K} x_{ijk} &= 1, & i \in I, j \in J, \\
x_{ijk} &= 1, & (i,j,k) \in A, \\
x_{ijk} &\in \{0,1\}, & i \in I, j \in J, k \in K.
\end{aligned}$$

(1.5p) b) Consider the objective function, to be minimized

$$f(x) := \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \bar{x}_{ijk} x_{ijk}.$$

Let $\tilde{\boldsymbol{x}} \in S$ and assume that $\tilde{\boldsymbol{x}} \neq \bar{\boldsymbol{x}}$. Let $\bar{k}_{ij}$ be the number assigned to tile (i, j) in solution $\bar{\boldsymbol{x}}$. Note that there exists by assumption at least one tile (i, j) such that $\tilde{x}_{ij\bar{k}_{ij}} = 0$. We yield that

$$f(\tilde{\boldsymbol{x}}) = \sum_{i \in I} \sum_{j \in J} \tilde{x}_{ij\bar{k}_{ij}} < \sum_{i \in I} \sum_{j \in J} 1 = \sum_{i \in I} \sum_{j \in J} \bar{x}_{ij\bar{k}_{ij}} \bar{x}_{ij\bar{k}_{ij}} = f(\bar{\boldsymbol{x}}).$$

Thus, $\bar{\boldsymbol{x}}$ is not an optimal solution. □

Question 6

(true or false)

(1p) a) True. By Weierstrass theorem, $f(\boldsymbol{y}) = \min_{\boldsymbol{x} \in S} ||\boldsymbol{y} - \boldsymbol{x}||$ has an optimal solution.

Suppose the optimal solution for $f(\boldsymbol{y}^1)$ is $\boldsymbol{x}^1$. For $f(\boldsymbol{y}^2)$ the optimal solution is $\boldsymbol{x}^2$.

$$\begin{aligned}
&\lambda f(\boldsymbol{y}^1) + (1-\lambda) f(\boldsymbol{y}^2)\\
=&\lambda \min_{\boldsymbol{x}\in S}\{||\boldsymbol{y}^1 - \boldsymbol{x}||\} + (1-\lambda)\min_{\boldsymbol{x}\in S}\{||\boldsymbol{y}^2 - \boldsymbol{x}||\}\\
=&\lambda||\boldsymbol{y}^1 - \boldsymbol{x}^1|| + (1-\lambda)||\boldsymbol{y}^2 - \boldsymbol{x}^2||\\
&\text{(by triangle-inequality)}\\
\geq&||\lambda(\boldsymbol{y}^1 - \boldsymbol{x}^1) + (1-\lambda)(\boldsymbol{y}^2 - \boldsymbol{x}^2)||\\
=&||\lambda\boldsymbol{y}^1 + (1-\lambda)\boldsymbol{y}^2 - (\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2)||\\
&\text{since } S \text{ is convex, } \boldsymbol{x}^1 \text{ and } \boldsymbol{x}^2 \text{ belong to } S,\ \lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \text{ also belong to } S\\
\geq&\min_{\boldsymbol{x}\in S}\{||[\lambda\boldsymbol{y}_1 + (1-\lambda)\boldsymbol{y}_2] - \boldsymbol{x}||\}\\
=&f(\lambda\boldsymbol{y}_1 + (1-\lambda)\boldsymbol{y}_2)
\end{aligned}$$

Thus, the function f is convex.

(1p) b) False. Suppose the feasible set is $x_1^2 + x_2 \leq 0,\ x_1^2 - x_2 \leq 0$, and the objective function (to be minimized) is $f = x_1$. Since the only feasible point is $(0,0)^T$, and the objective function is convex, the problem is convex. Thus, the KKT conditions are sufficient. But at point $(0,0)^T$, the gradient cone is $(a,0)^T$ where $a \in R$, and the tangent cone is $(0,0)^T$, so they are not the same. Thus, the KKT conditions are not necessary.

(1p) c) False. If no feasible solution exists, the optimal value is > 0. If feasible solutions exist, the optimal value is $= 0$.

(3p) Question 7

(Lagrangian relaxation and decomposition)

(1p) a) The Lagrangian dual function is

$$h(\boldsymbol{u}) = \inf\left\{\left(1 - \sum_{i\in\mathcal{I}} u_i\right) z + \sum_{i\in\mathcal{I}} u_i \sum_{j\in\mathcal{J}} p_{ij}x_{ij} \,\middle|\, \sum_{i\in I} x_{ij} = 1, j \in J, x_{ij} \in \mathbb{B}, z \in \mathbb{R}\right\}$$

Since there are no constraints on z we yield that $h(\boldsymbol{u}) = -\infty$ unless the coefficient $1 - \sum_{i\in\mathcal{I}} u_i$ is zero, i.e., $\sum_{i\in\mathcal{I}} u_i = 1$.

(1.5p) b) Note that there is no constraint that connects variables from different tasks and the objective is linear. By also assuming $\sum_{i\in\mathcal{I}} \bar{u}_i = 1$ we yield

$$h(\bar{\boldsymbol{u}}) = \sum_{j\in J} \min\left\{\sum_{i\in\mathcal{I}} \bar{u}_i p_{ij} x_{ij} \,\middle|\, \sum_{i\in I} x_{ij} = 1, x_{ij} \in \mathbb{B}, i \in I\right\}$$

The constraints can be read as choose one machine for each task, hence choosing a machine with (tied) smallest objective coefficient is optimal. Hence, let $i_j^* \in$

$\operatorname{argmin}_{i \in I} \bar{u}_i p_{ij}$, $j \in J$. The minimizer of the Lagrangian function at $\bar{u}$ is thus $\bar{x}_{i_j^* j} = 1$ for $j \in J$ and otherwise zero. We yield

$$h(\bar{\boldsymbol{u}}) = \sum_{j \in \mathcal{J}} \min_{i \in I} \bar{u}_i p_{ij}$$

(0.5p) c) All relaxed constraints are satisfied by choosing $\bar{z} = \operatorname{argmax}_{i \in I} \sum_{j \in J} p_{ij} \bar{x}_{ij}$, hence $(\bar{\boldsymbol{x}}, \bar{z})$ forms a primal feasible solution.