Solutions to exam in SF1811 Optimization, March 14, 2014

```
   1      3      5        1      3      5      1      3      5
30 0------0------0 -40    0      0------0      0      0------0
    \    / \    /          \    / \    /        \    /      /
     \  /   \  /            \  /   \  /          \  /      /
      \/     \/              \/     \/            \/      /
      /\     /\              /\     /\            /\     /
     /  \   /  \            /  \   /  \          /  \   /
    /    \ /    \          /    \ /    \        /    \ /
20 0------0------0 -10    0      0      0      0      0------0
   2      4      6        2      4      6      2      4      6
      Figure 1.          Figure 2.            Figure 3.
     The network        First span.tree      Second span.tree
```

1. Here, a mainly "algebraic" solution is presented. However, it is permitted, easier, and highly recommended, that all the calculations are made in figures of the network!

Let the two supply nodes be called "node 1" and "node 2", the two transhipment nodes "node 3" and "node 4", and the two demand nodes "node 5" and "node 6", see **Figure 1**. Then the set of arcs is given by $\mathcal{B} = \{(1,3),\,(1,4),\,(2,3),\,(2,4),\,(3,5),\,(3,6),\,(4,5),\,(4,6)\}$, and the minimum cost network flow problem can be written as the LP problem

$$\text{minimize } \mathbf{c}^\mathsf{T}\mathbf{v} \text{ subject to } \mathbf{A}\mathbf{v} = \mathbf{b},\ \mathbf{v} \geq \mathbf{0},\ \text{where}$$

$\mathbf{v} = (v_{13},\, v_{14},\, v_{23},\, v_{24},\, v_{35},\, v_{36},\, v_{45},\, v_{46})^\mathsf{T} = (x_{11},\, x_{12},\, x_{21},\, x_{22},\, z_{11},\, z_{12},\, z_{21},\, z_{22})^\mathsf{T}$,
$\mathbf{c} = (c_{13},\, c_{14},\, c_{23},\, c_{24},\, c_{35},\, c_{36},\, c_{45},\, c_{46})^\mathsf{T} = (p_{11},\, p_{12},\, p_{21},\, p_{22},\, q_{11},\, q_{12},\, q_{21},\, q_{22})^\mathsf{T}$,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \end{bmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 30 \\ 20 \\ 0 \\ 0 \\ -40 \end{pmatrix}.$$

The equation corresponding to node 6 has been removed since it is a linear combination of the other equations. (As always for balanced network flow problems.)

1.(a) The suggested solution corresponds to the spanning tree in **Figure 2**, consisting of the basic arcs $\mathcal{B}_\beta = \{(1,4),\,(2,3),\,(3,5),\,(3,6),\,(4,5)\}$, with the corresponding basis matrix

$$\mathbf{A}_\beta = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 \end{bmatrix}, \text{ which has linearly independent columns.}$$

The values of the corresponding basic variables, i.e. the flows in the basic arcs, can be determined as follows, which is equivalent to solving $\mathbf{A}_\beta \mathbf{v}_\beta = \mathbf{b}$:
$v_{14} = x_{12} = 30$, due to the flow balance equation in node 1,
$v_{23} = x_{21} = 20$, due to the flow balance equation in node 2,
$v_{45} = z_{21} = 30$, due to the flow balance equation in node 4,

$v_{35} = z_{11} = 10$, due to the flow balance equation in node 5,
$v_{36} = z_{12} = 10$, due to the flow balance equation in node 3,
which is the suggested solution! Since all $v_{ij} \geq 0$, it is a feasible basic solution.
The objective value is $\sum_{(i,j)} c_{ij} v_{ij} = 430$.

1.(b) To check if the current solution is optimal, the vector $\mathbf{y} = (y_1, \ldots, y_6)^\mathsf{T}$ of simplex multipliers is calculated from the equations $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$, which here become: $y_i - y_j = c_{ij}$ for all basic arcs, and $y_6 = 0$.
The basic arc $(3,6)$ implies that $y_3 - y_6 = c_{36}$, i.e. $y_3 = y_6 + c_{36} = 0 + 5 = 5$.
The basic arc $(3,5)$ implies that $y_3 - y_5 = c_{35}$, i.e. $y_5 = y_3 - c_{35} = 5 - 5 = 0$.
The basic arc $(4,5)$ implies that $y_4 - y_5 = c_{45}$, i.e. $y_4 = y_5 + c_{45} = 0 + 7 = 7$.
The basic arc $(2,3)$ implies that $y_2 - y_3 = c_{23}$, i.e. $y_2 = y_3 + c_{23} = 5 + 3 = 8$.
The basic arc $(1,4)$ implies that $y_1 - y_4 = c_{14}$, i.e. $y_1 = y_4 + c_{14} = 7 + 2 = 9$.

Next step is to calculate the reduced costs from $\mathbf{r}_\nu = \mathbf{c}_\nu - \mathbf{A}_\nu^\mathsf{T} \mathbf{y}$, which here becomes: $r_{ij} = c_{ij} - y_i + y_j$ for all nonbasic arcs.
$r_{13} = c_{13} - y_1 + y_3 = 5 - 9 + 5 = 1$,
$r_{24} = c_{24} - y_2 + y_4 = 2 - 8 + 7 = 1$
$r_{46} = c_{46} - y_4 + y_6 = 6 - 7 + 0 = -1$.

Since $r_{46} < 0$, the non-basic variable v_{46} should become a new basic variable.
Let $v_{46} = t$ and let t increase from zero.
Then the basic variables, i.e. the flows in the basic arc, are affected as follows:
$v_{14} = x_{12} = 30$, due to the flow balance equation in node 1.
$v_{23} = x_{21} = 20$, due to the flow balance equation in node 2.
$v_{45} = z_{21} = 30 - t$, due to the flow balance equation in node 4.
$v_{35} = z_{11} = 10 + t$, due to the flow balance equation in node 5.
$v_{36} = z_{12} = 10 - t$, due to the flow balance equation in node 3.

We see that t can be increased to $t = 10$.
Then v_{36} has decreased to zero and should be replaced by v_{46} as basic variable.
The new feasible basic solution, corresponding to the spanning tree in **Figure 3**, is
$v_{14} = x_{12} = 30$, $v_{23} = x_{21} = 20$, $v_{45} = z_{21} = 20$, $v_{35} = z_{11} = 20$, $v_{46} = z_{22} = 10$,
$v_{13} = x_{11} = 0$, $v_{24} = x_{22} = 0$, $v_{36} = z_{12} = 0$, with objective value $\sum_{(i,j)} c_{ij} v_{ij} = 420$.

The corresponding vector $\mathbf{y} = (y_1, \ldots, y_6)^\mathsf{T}$ is calculated from the equations
$y_i - y_j = c_{ij}$ for all basic arcs, and $y_6 = 0$.
The basic arc $(4,6)$ implies that $y_4 - y_6 = c_{46}$, i.e. $y_4 = y_6 + c_{46} = 0 + 6 = 6$.
The basic arc $(4,5)$ implies that $y_4 - y_5 = c_{45}$, i.e. $y_5 = y_4 - c_{45} = 6 - 7 = -1$.
The basic arc $(3,5)$ implies that $y_3 - y_5 = c_{35}$, i.e. $y_3 = y_5 + c_{35} = -1 + 5 = 4$.
The basic arc $(2,3)$ implies that $y_2 - y_3 = c_{23}$, i.e. $y_2 = y_3 + c_{23} = 4 + 5 = 7$.
The basic arc $(1,4)$ implies that $y_1 - y_4 = c_{14}$, i.e. $y_1 = y_4 + c_{14} = 6 + 2 = 8$.

The corresponding reduced costs are calculated from the formula
$r_{ij} = c_{ij} - y_i + y_j$ for all non-basic arcs.
$r_{13} = c_{13} - y_1 + y_3 = 5 - 8 + 4 = 1$,
$r_{24} = c_{24} - y_2 + y_4 = 2 - 7 + 6 = 1$
$r_{36} = c_{36} - y_3 + y_6 = 5 - 4 + 0 = 1$.

Since all $r_{ij} \geq 0$, the current solution is optimal!

2.(a). The considered LP problem is on the form: minimize $\mathbf{c}^\mathsf{T}\mathbf{z}$ subject to $\mathbf{G}\mathbf{z} = \mathbf{b},\ \mathbf{z} \geq \mathbf{0}$, where $\mathbf{z} = (x_1, x_2, x_3, x_4, v_1, v_2, v_3)^\mathsf{T}$, $\mathbf{c} = (0,\ 0,\ 0,\ 0,\ 1,\ 1,\ 1)^\mathsf{T}$,

$$\mathbf{G} = [\mathbf{A}\ \ \mathbf{I}] = \begin{bmatrix} 1 & 0 & -1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}.$$

The suggested solution corresponds to a basis with $\beta = (1, 2, 7)$. The values of the basic variables in this basic solution are obtained from the system $\mathbf{G}_\beta \mathbf{z}_\beta = \mathbf{b}$, i.e.

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}, \text{ with the unique solution } \begin{pmatrix} x_1 \\ x_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}, \quad \textbf{OK!}$$

The corresponding simplex multipliers are obtained from the system $\mathbf{G}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$, i.e.

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{ with the unique solution } \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

The reduced costs for the non-basic variables are obtained from

$$\mathbf{r}_\delta^\mathsf{T} = \mathbf{c}_\delta - \mathbf{y}^\mathsf{T}\mathbf{G}_\delta = (\,0,\ 0,\ 1,\ 1\,) - (-1, -1,\ 1\,) \begin{bmatrix} -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 1 & 0 & 0 \end{bmatrix} = (\,0,\ 1,\ 2,\ 2\,).$$

Since $\mathbf{r}_\delta \geq \mathbf{0}$, the suggested feasible basic solution is optimal.
The optimal value of the problem is $\mathbf{c}^\mathsf{T}\mathbf{z} = 1$.

2.(b). The answer is NO, because of the following arguments:

Assume that there were scalars $x_j \geq 0$ such that $\mathbf{b} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 + \mathbf{a}_4 x_4$, so that the vector $\mathbf{x} = (x_1, x_2, x_3, x_4)^\mathsf{T}$ would satisfy $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.
Then this vector $\mathbf{x}$ together with the vector $\mathbf{v} = \mathbf{0}$ would be a feasible solution to the above LP problem with objective value $= v_1 + v_2 + v_3 = 0 + 0 + 0 = 0$. But this is a contradiction, since we already know that the optimal value of the LP problem is $= 1$.

2.(c). The dual problem corresponding to the above "primal" LP problem can be written

$$\begin{aligned} \text{maximize} \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{subject to} \quad & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{0}, \\ & \mathbf{I}\,\mathbf{y} \leq \mathbf{e}. \end{aligned}$$

The Duality Theorem for LP implies that since the primal LP problem has an *optimal* solution with objective value $=1$ ($=$ the optimal value), the dual problem will also have an optimal solution with (dual) objective value $=1$.
Any such optimal solution $\mathbf{y}$ to the dual problem thus satisfies $\mathbf{b}^\mathsf{T}\mathbf{y} = 1$, $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{0}$ and $\mathbf{I}\,\mathbf{y} \leq \mathbf{e}$, and thus, in particular, $\mathbf{b}^\mathsf{T}\mathbf{y} > 0$ and $\mathbf{a}_j^\mathsf{T}\mathbf{y} \leq 0$ for all j.
But an optimal solution to the dual problem is given by the vector of simplex multipliers corresponding to an optimal basis for the primal problem, in our case $\mathbf{y} = (-1, -1,\ 1\,)^\mathsf{T}$.
This vector satisfies $\mathbf{b}^\mathsf{T}\mathbf{y} = 1 > 0$, $\mathbf{a}_1^\mathsf{T}\mathbf{y} = 0$, $\mathbf{a}_2^\mathsf{T}\mathbf{y} = 0$, $\mathbf{a}_3^\mathsf{T}\mathbf{y} = 0$ and $\mathbf{a}_4^\mathsf{T}\mathbf{y} = -1 \leq 0$.

3.(a) Since $\frac{1}{2}\|\mathbf{x}-\mathbf{q}\|^2 = \frac{1}{2}(\mathbf{x}-\mathbf{q})^\mathsf{T}(\mathbf{x}-\mathbf{q}) = \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{I}\,\mathbf{x} - \mathbf{q}^\mathsf{T}\mathbf{x} + \frac{1}{2}\,\mathbf{q}^\mathsf{T}\mathbf{q}$, the considered problem can be written: minimize $\frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

$$\mathbf{H} = \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\quad \mathbf{c} = -\mathbf{q} = \begin{pmatrix} -q_1 \\ -q_2 \\ -q_3 \\ -q_4 \end{pmatrix},\quad \mathbf{A} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix},\quad \mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The matrix $\mathbf{H} = \mathbf{I}$ is positive definite, so we have a convex QP problem for which the following Lagrange optimality conditions are both necessary and sufficient conditions for a global optimum:

$$\begin{aligned} \mathbf{H}\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{u} &= -\mathbf{c} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \end{aligned}$$

The equations $\mathbf{H}\mathbf{x} - \mathbf{A}^\mathsf{T}\mathbf{u} = -\mathbf{c}$ are in our case equivalent to $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{u} + \mathbf{q}$.

If this is combined with the remaining equations $\mathbf{A}\mathbf{x} = \mathbf{0}$, we get that $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{u} = -\mathbf{A}\mathbf{q}$,

which in our case becomes $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -q_1 - q_2 + q_3 + q_4 \\ -q_1 + q_2 - q_3 + q_4 \end{pmatrix}$,

with the unique solution $\hat{\mathbf{u}} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \dfrac{1}{4}\begin{pmatrix} -q_1 - q_2 + q_3 + q_4 \\ -q_1 + q_2 - q_3 + q_4 \end{pmatrix}$.

The corresponding unique $\hat{\mathbf{x}}$, which together with $\hat{\mathbf{u}}$ satisfies the Lagrange conditions,

is then given by $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\hat{\mathbf{u}} + \mathbf{q} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}\begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} + \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} = \dfrac{1}{2}\begin{pmatrix} q_1 + q_4 \\ q_2 + q_3 \\ q_3 + q_2 \\ q_4 + q_1 \end{pmatrix}$.

3.(b) With $\mathbf{y} = \mathbf{A}^\mathsf{T}\mathbf{v}$, the objective function becomes

$\frac{1}{2}\|\mathbf{A}^\mathsf{T}\mathbf{v} - \mathbf{q}\|^2 = \frac{1}{2}(\mathbf{A}^\mathsf{T}\mathbf{v} - \mathbf{q})^\mathsf{T}(\mathbf{A}^\mathsf{T}\mathbf{v} - \mathbf{q}) = \frac{1}{2}\mathbf{v}^\mathsf{T}\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{v} - \mathbf{q}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{v} + \frac{1}{2}\mathbf{q}^\mathsf{T}\mathbf{q}$.

Since $\mathbf{A}\mathbf{A}^\mathsf{T} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ is positive definite, the necessary and sufficient optimality conditions for the considered minimization problem are given by the equations $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{v} = \mathbf{A}\mathbf{q}$,

which in our case becomes $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} q_1 + q_2 - q_3 - q_4 \\ q_1 - q_2 + q_3 - q_4 \end{pmatrix}$,

with the unique solution $\hat{\mathbf{v}} = \begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \end{pmatrix} = \dfrac{1}{4}\begin{pmatrix} q_1 + q_2 - q_3 - q_4 \\ q_1 - q_2 + q_3 - q_4 \end{pmatrix}$.

The corresponding unique $\hat{\mathbf{y}}$ is then given by $\hat{\mathbf{y}} = \mathbf{A}^\mathsf{T}\hat{\mathbf{v}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}\begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \end{pmatrix} = \dfrac{1}{2}\begin{pmatrix} q_1 - q_4 \\ q_2 - q_3 \\ q_3 - q_2 \\ q_4 - q_1 \end{pmatrix}$.

We note that $\hat{\mathbf{y}}^\mathsf{T}\hat{\mathbf{x}} = 0$ and $\hat{\mathbf{x}} + \hat{\mathbf{y}} = \mathbf{q}$, as it should be.

4.(a)
The objective function is $f(\mathbf{x}) = x_1^2x_2^2 + x_1^2 + 3x_2^2 - 2x_1x_2 - 4x_1 - 4x_2$.

The gradient of f becomes $\nabla f(\mathbf{x}) = (2x_1x_2^2 + 2x_1 - 2x_2 - 4,\ 2x_1^2x_2 + 6x_2 - 2x_1 - 4)$.

The Hessian of f becomes $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 2x_2^2+2 & 4x_1x_2-2 \\ 4x_1x_2-2 & 2x_1^2+6 \end{bmatrix}$.

The starting point is given by $\mathbf{x}^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, with $f(\mathbf{x}^{(1)}) = 0$.

$\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix}$ is positive definite since $2 > 0$, $6 > 0$ and $2 \cdot 6 - (-2)\cdot(-2) > 0$.

Then the first Newton search direction $\mathbf{d}^{(1)}$ is obtained by solving the system
$\mathbf{F}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla f(\mathbf{x}^{(1)})^\mathsf{T}$, i.e. $\begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix}\mathbf{d} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, with the solution $\mathbf{d}^{(1)} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

First try $t_1 = 1$, so that $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

Then $f(\mathbf{x}^{(2)}) = 52 > f(\mathbf{x}^{(1)})$, so $t_1 = 1$ is not accepted.

Then try $t_1 = 0.5$, so that $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + 0.5\mathbf{d}^{(1)} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Then $f(\mathbf{x}^{(2)}) = -5 < f(\mathbf{x}^{(1)})$, so $t_1 = 0.5$ is accepted, and the first iteration is completed.

4.(b)
f is convex on $\mathbb{R}^2$ if and only if the Hessian $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 2x_2^2+2 & 4x_1x_2-2 \\ 4x_1x_2-2 & 2x_1^2+6 \end{bmatrix}$
is positive semidefinite for all $x \in \mathbb{R}^2$.

Both the diagonal elements in $\mathbf{F}(\mathbf{x})$ are clearly always > 0, so $\mathbf{F}(\mathbf{x})$ is is positive semidefinit if and only if $(2x_2^2 + 2)(2x_1^2 + 6) - (4x_1x_2 - 2)(4x_1x_2 - 2) \geq 0$.

But $(2x_2^2 + 2)(2x_1^2 + 6) - (4x_1x_2 - 2)(4x_1x_2 - 2) = 8 + 4x_1^2 + 12x_2^2 + 16x_1x_2 - 12x_1^2x_2^2$,
which becomes negative if both x_1 and x_2 are sufficiently large numbers:
If $x_1 = x_2 = 10$ then $8 + 4x_1^2 + 12x_2^2 + 16x_1x_2 - 12x_1^2x_2^2 = 8 + 400 + 1200 + 1600 - 120000 < 0$.
The conclusion is that f is not convex on $\mathbb{R}^2$.

4.(c)
Now $f(\mathbf{x})$ should be minimized subject to the constraint $x_1 - x_2 = 0$.

We will use a nullspace approach: The complete set of feasible solutions to the constraint is given by $\mathbf{x}(t) = \begin{pmatrix} t \\ t \end{pmatrix}$, with $t \in \mathbb{R}$. If this is plugged into the objective function, it becomes:
$h(t) = f(\mathbf{x}(t)) = t^4 + 2t^2 - 8t$, with $h'(t) = 4t^3 + 4t - 8$ and $h''(t) = 12t^2 + 4 > 0$ for all t.

Thus $h(t)$ is strictly convex, so if we find a $t_0 \in \mathbb{R}$ such that $h'(t_0) = 0$ then t_0 must be the unique optimal solution to the problem of minimizing $h(t)$.

But we immediatly see that $t_0 = 1$ is such a solution, which means that $\hat{\mathbf{x}} = (1, 1)^\mathsf{T}$ is a globally optimal solution to the problem of minimizing $f(\mathbf{x})$ subject to $x_1 - x_2 = 0$.

5. The feasible region of P is a convex set, since all the constraints are linear.

Further, the objective function $f(\mathbf{x}) = \sum_{i=1}^{2}\sum_{j=1}^{3}(x_{ij}\ln(x_{ij}) - x_{ij})$ may be written

$f(\mathbf{x}) = \sum_{i=1}^{2}\sum_{j=1}^{3} f_{ij}(x_{ij})$, where $f_{ij}(x_{ij}) = x_{ij}\ln(x_{ij}) - x_{ij}$.

Some calculus give that $f'_{ij}(x_{ij}) = \ln(x_{ij})$ and $f''_{ij}(x_{ij}) = 1/x_{ij} > 0$, which implies that the Hessian matrix $\mathbf{F}(\mathbf{x})$ is a diagonal matrix with diagonal elements $1/x_{ij} > 0$ for all $\mathbf{x} \in X$, which in turn implies that f is a *strictly convex* function on X.

5.(a) Let $a_1 = 2/5$, $a_2 = 3/5$, $b_1 = 1/6$, $b_2 = 1/3$ and $b_3 = 1/2$.

Further, let (λ, μ) be a shorter notation for $(\lambda_1, \lambda_2, \mu_1, \mu_2, \mu_3)$.

Then the Lagrange function for the considered problem is given by

$$L(\mathbf{x}, \lambda, \mu) = \sum_{i=1}^{2}\sum_{j=1}^{3}(x_{ij}\ln(x_{ij}) - x_{ij}) + \lambda_1(x_{11} + x_{12} + x_{13} - a_1) + \lambda_2(x_{21} + x_{22} + x_{23} - a_2) +$$
$$+ \mu_1(x_{11} + x_{21} - b_1) + \mu_2(x_{12} + x_{22} - b_2) + \mu_3(x_{13} + x_{23} - b_3) =$$
$$\sum_{i=1}^{2}\sum_{j=1}^{3}(x_{ij}\ln(x_{ij}) - x_{ij} + \lambda_i x_{ij} + \mu_j x_{ij}) - \sum_{i=1}^{2} a_i\lambda_i - \sum_{j=1}^{3} b_j\mu_j.$$

The Lagrange relaxed problem $\mathrm{PR}_{\lambda,\mu}$ is defined, for a given vector $(\lambda, \mu)^{\mathsf{T}} \geq \mathbf{0}$, as the problem of minimizing $L(\mathbf{x}, \lambda, \mu)$ with respect to $\mathbf{x} > \mathbf{0}$.
But this problem separates into one problem for each variables x_{ij}, namely the problem of minimizing $\ell_{ij}(x_{ij}) = x_{ij}\ln(x_{ij}) - x_{ij} + \lambda_i x_{ij} + \mu_j x_{ij}$ subject to $x_{ij} > 0$.

Some calculus give that $\ell'_{ij}(x_{ij}) = \ln(x_{ij}) + \lambda_i + \mu_j$ and $\ell''_{ij}(x_{ij}) = \dfrac{1}{x_{ij}} > 0$,

which implies that $\ell_{ij}(x_{ij})$ is strictly convex on the set $x_{ij} > 0$, and a minimizing x_{ij} is obtained from the equation $\ell'_{ij}(x_{ij}) = 0$, which has the unique solution

$\tilde{x}_{ij}(\lambda, \mu) = e^{-\lambda_i - \mu_j}$, and then the dual objective function becomes

$$\varphi(\lambda, \mu) = L(\tilde{\mathbf{x}}(\lambda, \mu), \lambda, \mu) = -\sum_{i=1}^{2}\sum_{j=1}^{3} e^{-\lambda_i - \mu_j} - \sum_{i=1}^{2} a_i\lambda_i - \sum_{j=1}^{3} b_j\mu_j.$$

The dual problem consists of maximizing $\varphi(\lambda, \mu)$ with respect to $(\lambda, \mu)^{\mathsf{T}} \geq \mathbf{0}$.

5.(b) The following solution is suggested: (note that $\ln(5/2) = -\ln(a_1)$, etc.)

$(\hat{\lambda}, \hat{\mu}) = (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3) = -(\ln(a_1), \ln(a_2), \ln(b_1), \ln(b_2), \ln(b_3))$.

We should try to find a corresponding vector $\hat{\mathbf{x}}$ such that $\hat{\mathbf{x}}$ and $(\hat{\lambda}, \hat{\mu})$ satisfy the global optimality conditions (GOC).

In order to satisfy the *first* GOC condition, i.e. $L(\hat{\mathbf{x}}, \hat{\lambda}, \hat{\mu}) \leq L(\mathbf{x}, \hat{\lambda}, \hat{\mu})$ for all $\mathbf{x} \in X$, $\hat{\mathbf{x}}$ must be chosen according to the formula
$\hat{x}_{ij} = \tilde{x}_{ij}(\hat{\lambda}, \hat{\mu}) = e^{-\hat{\lambda}_i - \hat{\mu}_j} = e^{-\hat{\lambda}_i} e^{-\hat{\mu}_j} = a_i b_j$, i.e.

$\hat{\mathbf{x}} = (\hat{x}_{11}, \hat{x}_{12}, \hat{x}_{13}, \hat{x}_{21}, \hat{x}_{22}, \hat{x}_{23}) = (2/30,\ 4/30,\ 6/30,\ 3/30,\ 6/30,\ 9/30)$.

Then the five explicit constraints in P are satisfied with *equality* ($\hat{x}_{11}+\hat{x}_{12}+\hat{x}_{13}=2/5$, etc.) so that both the *second* and the *fourth* GOC are satisfied. Finally, the *third* GOC is satisfied since $\hat{\lambda}_i = -\ln(a_i) = \ln(1/a_i) > 0$ and $\hat{\mu}_j = -\ln(b_j) = \ln(1/b_j) > 0$ for all i and j.
Thus $\hat{\mathbf{x}}$ is a global optimal solution to P and $(\hat{\lambda}, \hat{\mu})$ is a global optimal solution to D.

As a check, the primal optimal value is equalt to $f(\hat{\mathbf{x}}) = \sum_{i=1}^{2}\sum_{j=1}^{3}(\hat{x}_{ij}\ln(\hat{x}_{ij}) - \hat{x}_{ij})$,

while the dual optimal value is $\varphi(\hat{\lambda},\hat{\mu}) = L(\tilde{\mathbf{x}}(\hat{\lambda},\hat{\mu}),\hat{\lambda},\hat{\mu}) = L(\hat{\mathbf{x}},\hat{\lambda},\hat{\mu}) =$
$f(\hat{\mathbf{x}}) + \hat{\lambda}_1(\hat{x}_{11}+\hat{x}_{12}+\hat{x}_{13}-a_1) + \hat{\lambda}_2(\hat{x}_{21}+\hat{x}_{22}+\hat{x}_{23}-a_2) +$
$\hat{\mu}_1(\hat{x}_{11}+\hat{x}_{21}-b_1) + \hat{\mu}_2(\hat{x}_{12}+\hat{x}_{22}-b_2) + \hat{\mu}_3(\hat{x}_{13}+\hat{x}_{23}-b_3) = f(\hat{\mathbf{x}})$,
since, according to above, all the other terms are zero.

5.(c)

As pointed out above, $f(\mathbf{x})$ is a strictly convex function on the convex feasible region of P.

We know from 5.(b) that $\hat{\mathbf{x}}$ defined by $\hat{x}_{ij} = a_i b_j$ is an optimal solution to P.
Assume that there is also another optimal solution $\bar{\mathbf{x}} \neq \hat{\mathbf{x}}$ (which thus satisfies $f(\bar{\mathbf{x}}) = f(\hat{\mathbf{x}})$).
Then (due to the convexity of the feasible region) $\frac{1}{2}(\hat{\mathbf{x}}+\bar{\mathbf{x}})$ is also a feasible solution to P, and (due to the strict convexity of f)

$f(\frac{1}{2}(\hat{\mathbf{x}}+\bar{\mathbf{x}})) < \frac{1}{2}f(\hat{\mathbf{x}}) + \frac{1}{2}f(\bar{\mathbf{x}}) = f(\hat{\mathbf{x}})$, which is a contradiction.

Thus, $\hat{\mathbf{x}}$ is the unique optimal solution to P.

We also know from 5.(b) that $(\hat{\lambda},\hat{\mu}) = -(\ln(a_1), \ln(a_2), \ln(b_1), \ln(b_2), \ln(b_3))$
is an optimal solution to D. In particular, this optimal dual solution satisfies

$\tilde{x}_{ij}(\hat{\lambda},\hat{\mu}) = e^{-\hat{\lambda}_i - \hat{\mu}_j} = e^{\ln a_i + \ln b_j} = a_i b_j = \hat{x}_{ij}$ for all i and j.

The question is if there is some other dual solution $(\lambda,\mu)^{\mathsf{T}} \geq \mathbf{0}$ such that

$\tilde{x}_{ij}(\lambda,\mu) = e^{-\lambda_i - \mu_j} = a_i b_j = \hat{x}_{ij}$ for all i and j.

The answer is yes! By letting

$(\lambda,\mu) = (\lambda_1,\lambda_2,\mu_1,\mu_2,\mu_3) = -(\ln(a_1)+c,\ \ln(a_2)+c,\ \ln(b_1)-c,\ \ln(b_2)-c,\ \ln(b_3)-c)$

for some constant c, it follows that

$\tilde{x}_{ij}(\lambda,\mu) = e^{-\lambda_i - \mu_j} = e^{\ln a_i + c + \ln b_j - c} = a_i b_j = \hat{x}_{ij}$.

This dual solution satisfies $(\lambda,\mu)^{\mathsf{T}} \geq \mathbf{0}$ if and only if $\ln(b_j) \leq c \leq -\ln(a_i)$ for all i and j, which in our case is satisfied if and only if $\ln(1/2) \leq c \leq \ln(5/3)$.

Since $\ln(1/2) < 0$ and $\ln(5/3) > 0$, there are such constants, e.g. $c = \ln(5/6)$,

in which case $(\lambda,\mu) = (\lambda_1,\lambda_2,\mu_1,\mu_2,\mu_3) = (\ln(3), \ln(2), \ln(5), \ln(5/2), \ln(5/3))$.

This dual solution, together with $\hat{\mathbf{x}}$, also satisfies the global optimality conditions,
and it is therefore also an optimal solution to D.