# Solutions to exam in SF1811 Optimization, Jan 11, 2017

**1.(a)** The network is illustrated in the Figure 1 below, where the supply at the nodes are written in the figure. Negative supply means demand. All arcs are directed from left to right. Since it is a balances problem (total demand equals total supply), and since the planner's suggestion corresponds to a spanning tree in the network, the values of the corresponding basic variables, i.e. the flows in the spanning tree arcs, can be determined as follows:
$x_{14} = 40$, due to the flow balance requirement in node 1,
$x_{23} = 30$, due to the flow balance requirement in node 2,
$x_{46} = 40$, due to the flow balance requirement in node 4,
$x_{35} = 20$, due to the flow balance requirement in node 5,
$x_{36} = 10$, due to the flow balance requirement in node 3.
These values are written at the arcs in Figure 2 below.

```
   1      3      5            1      3      5
40 O------O------O -20     40 O      O--20--O -20
    \    / \    /              \    / \
     \  /   \  /               40  /   \
      \/     \/                  \/     10
      /\     /\                  /\       \
     /  \   /  \               30  \       \
    /    \ /    \              /    \       \
30 O------O------O -50     30 O      O--40--O -50
   2      4      6            2      4      6
  Figure 1. The network.     Figure 2. The basic solution.
```

Since the solution correspond to a spanning tree, the flow is balances in every node, and every $x_{ij} \geq 0$, it is a BFS (*basic feasible solution*). To check if it is an optimal solution, we use the simplex method for network flow problems, starting from this BFS.
First, the simplex multipliers $y_i$ for the nodes are calculated from $y_6 = 0$ and $y_i - y_j = c_{ij}$ for all ars $(i, j)$ in the spanning tree, see Figure 3 below where the costs $c_{ij}$ for arcs in the spanning tree are written at the arcs.
Then the reduced cost for the non-basic variables are calculated from $r_{ij} = c_{ij} - y_i + y_j$, see Figure 4 below, where the costs $c_{ij}$ for arcs *not* in the spanning tree are written at the arcs.

```
     y1=5   y3=4   y5=1              y1=5   y3=4   y5=1
      O      O--3---O                  O--2---O      O
       \    / \                     r13=2-5+4=1     /
        2  /   \                                   /
         \/     4                                 3  r45=3-3+1=1
         /\      \                               /
        2  \      \                             /
       /    \      \                r24=4-6+3=1/
      O      O--3---O                  O--4---O      O
     y2=6   y4=3   y6=0              y2=6   y4=3   y6=0
  Figure 3. Calculations of y_i.  Figure 4. Calculations of r_{ij}.
```

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

**1.(b)** The minimum cost network flow problem can be written as the LP problem

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

where $\mathbf{x} = (x_{13},\, x_{14},\, x_{23},\, x_{24},\, x_{35},\, x_{36},\, x_{45},\, x_{46})^\mathsf{T}$,

$\mathbf{c} = (c_{13},\, c_{14},\, c_{23},\, c_{24},\, c_{35},\, c_{36},\, c_{45},\, c_{46})^\mathsf{T} = (2, 2, 2, 4, 3, 4, 3, 3)^\mathsf{T}$,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 40 \\ 30 \\ 0 \\ 0 \\ -20 \\ -50 \end{pmatrix}.$$

(The equation corresponding to node 6 can be removed since it is a linear combination of the other five equations, but that is not necessary and has not been done here.)

When the primal problem is on the above standard form, the corresponding dual problem is

$$\text{maximize } \mathbf{b}^\mathsf{T}\mathbf{y} \text{ subject to } \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}\,,$$

which here becomes

$$\begin{aligned} \text{maximize} \quad & 40y_1 + 30y_2 + 0y_3 + 0y_4 - 20y_5 - 50y_6 \\ \text{subject to} \quad & y_1 - y_3 \leq 2, \\ & y_1 - y_4 \leq 2, \\ & y_2 - y_3 \leq 2, \\ & y_2 - y_4 \leq 4, \\ & y_3 - y_5 \leq 3, \\ & y_3 - y_6 \leq 4, \\ & y_4 - y_5 \leq 3, \\ & y_4 - y_6 \leq 3, \end{aligned}$$

It is well known that an optimal solution to this dual problem is given by the vector $\mathbf{y}$ with simplex multipliers from 1.(a), i.e. $\mathbf{y} = (5, 6, 4, 3, 1, 0)^\mathsf{T}$. Then the right hand sides minus the left hand sides of the dual constraint become $\mathbf{c} - \mathbf{A}^\mathsf{T}\mathbf{y} = (1,\, 0,\, 0,\, 1,\, 0,\, 0,\, 1,\, 0)^\mathsf{T} \geq \mathbf{0}$, which shows that $\mathbf{y}$ is a feasible solution to the dual problem.
The (dual) objective value of this solution is $\mathbf{b}^\mathsf{T}\mathbf{y} = 40 \cdot 5 + 30 \cdot 6 - 20 \cdot 1 - 50 \cdot 0 = 360$, while the (primal) objective value of the solution $\mathbf{x} = (0, 40, 30, 0, 20, 10, 0, 40)^\mathsf{T}$, calculated in (a) above, is $\mathbf{c}^\mathsf{T}\mathbf{x} = 2 \cdot 40 + 2 \cdot 30 + 3 \cdot 20 + 4 \cdot 10 + 3 \cdot 40 = 360$.

Since $\mathbf{x}$ is a feasible solution to the primal problem, $\mathbf{y}$ is a feasible solution to the dual problem, and $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{b}^\mathsf{T}\mathbf{y}$, the duality theorem tells us that $\mathbf{x}$ and $\mathbf{y}$ are optimal solution to, respectively, the primal and the dual problems.

**2.(a)** We have an LP problem on the standard form

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

where $\mathbf{A} = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$ and $\mathbf{c}^{\mathsf{T}} = (3,\ 3,\ 6)$.

If $x_2$ and $x_3$ are chosen as basic variables then $\beta = (2,3)$ and $\nu = (1)$,
and the corresponding basic matrix is $\mathbf{A}_\beta = \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}$, while $\mathbf{A}_\nu = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.
Since $\det(\mathbf{A}_\beta) \neq 0$, the columns of $\mathbf{A}_\beta$ are linearly independent, as they should be for a basic matrix, and then the values of the current basic variables are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is calculated from the system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e.

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1.5 \end{pmatrix}.$$

Since $\bar{\mathbf{b}} \geq \mathbf{0}$, the solution $\mathbf{x}_\beta = \bar{\mathbf{b}}$ and $\mathbf{x}_\nu = \mathbf{0}$ (i.e. $x_1 = 0,\ x_2 = 3,\ x_3 = 1.5$) is a BFS.

The vector $\mathbf{y}$ with simplex multipliers is obtained from the system $\mathbf{A}_\beta^{\mathsf{T}} \mathbf{y} = \mathbf{c}_\beta$, i.e.

$$\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \text{ with the solution } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Then the reduced cost for the non-basic variable is obtained from

$$\mathbf{r}_\nu^{\mathsf{T}} = \mathbf{c}_\nu^{\mathsf{T}} - \mathbf{y}^{\mathsf{T}} \mathbf{A}_\nu = 3 - (3,\ 0) \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -3.$$

Since $r_{\nu_1} = r_1 = -3 < 0$, we let $x_1$ become the new basic variable.

Then we should calculate the vector $\bar{\mathbf{a}}_1$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_1 = \mathbf{a}_1$, i.e.

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{pmatrix} \bar{a}_{11} \\ \bar{a}_{21} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{a}}_1 = \begin{pmatrix} \bar{a}_{11} \\ \bar{a}_{21} \end{pmatrix} = \begin{pmatrix} 3 \\ -0.5 \end{pmatrix}.$$

The largest permitted value of the new basic variable $x_1$ is then given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i1}} \mid \bar{a}_{i1} > 0 \right\} = \min \left\{ \frac{3}{3},\ - \right\} = \frac{3}{3} = \frac{\bar{b}_1}{\bar{a}_{11}}.$$

Minimizing index is $i = 1$, which implies that $x_{\beta_1} = x_2$ should no longer be a basic variable. Its place as basic variable is taken by $x_1$, so that $\beta = (1,3)$ and $\nu = (2)$.

The corresponding basic matrix is $\mathbf{A}_\beta = \begin{bmatrix} 2 & 2 \\ 1 & 4 \end{bmatrix}$, while $\mathbf{A}_\nu = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

The values of the current basic variables are $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is calculated from the system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e.

$$\begin{bmatrix} 2 & 2 \\ 1 & 4 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

The vector $\mathbf{y}$ with simplex multipliers is obtained from the system $\mathbf{A}_\beta^{\mathsf{T}} \mathbf{y} = \mathbf{c}_\beta$, i.e.

$$\begin{bmatrix} 2 & 1 \\ 2 & 4 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \text{ with the solution } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then the reduced cost for the non-basic variable is obtained from

$$\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\nu = 3 - (1,\,1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1.$$

Since $\mathbf{r}_\nu \geq \mathbf{0}$ the current feasible basic solution is optimal.
Thus, $\mathbf{x} = (1, 0, 2)^\mathsf{T}$ is an optimal solution, with optimal value $\mathbf{c}^\mathsf{T}\mathbf{x} = 15$.

**2.(b)** Now it is assumed that $\mathbf{b} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$ (instead of $\begin{pmatrix} 6 \\ 9 \end{pmatrix}$).

With $x_2$ and $x_3$ as basic variables, the values of these variables are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is calculated from the system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e.

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}\begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2.5 \end{pmatrix}.$$

Since $\bar{b}_1 < 0$, the basic solution $\mathbf{x}_\beta = \bar{\mathbf{b}}$ and $\mathbf{x}_\nu = \mathbf{0}$ (i.e. $x_1 = 0,\ x_2 = -1,\ x_3 = 2.5$) is NOT a BFS.

**2.(c)** Let $x_4 = v_1$ and $x_5 = v_2$.

Then the "Phase 1 problem" also becomes a problem on the standard form:

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

But now $\mathbf{A} = \begin{bmatrix} 2 & 1 & 2 & 1 & 0 \\ 1 & 1 & 4 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$ and $\mathbf{c}^\mathsf{T} = (0,\,0,\,0,\,1,\,1)$.

The suggested starting solution should have the basic variables $v_1$ and $v_2$, i.e. $x_4$ and $x_5$, which means that $\beta = (4, 5)$ and $\nu = (1, 2, 3)$.
The corresponding basic matrix is $\mathbf{A}_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, while $\mathbf{A}_\nu = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 4 \end{bmatrix}$.

The columns of $\mathbf{A}_\beta$ are linearly independent, as they should be for a basic matrix.
Then the values of the current basic variables are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$,
where the vector $\bar{\mathbf{b}}$ is calculated from the system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}.$$

Since $\bar{\mathbf{b}} \geq \mathbf{0}$, the solution $\mathbf{x}_\beta = \bar{\mathbf{b}}$ and $\mathbf{x}_\nu = \mathbf{0}$ (i.e. $x_1 = x_2 = x_3 = 0$, $v_1 = x_4 = 4$ and $v_2 = x_5 = 9$) is a BFS.

The vector $\mathbf{y}$ with simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta$, i.e.

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

Then the reduced cost for the non-basic variable is obtained from

$$\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\nu = (0,\,0,\,0) - (1,\,1)\begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 4 \end{bmatrix} = (-3,\,-2,\,-6).$$

Since $r_{\nu_3} = r_3 = -6$ is smallest, and $< 0$, we let $x_3$ become the new basic variable.

Then we should calculate the vector $\bar{\mathbf{a}}_3$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_3 = \mathbf{a}_3$, i.e.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{a}}_3 = \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

The largest permitted value of the new basic variable $x_3$ is then given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \min \left\{ \frac{4}{2}, \frac{9}{4} \right\} = \frac{4}{2} = \frac{\bar{b}_1}{\bar{a}_{13}}.$$

Minimizing index is $i = 1$, which implies that $x_{\beta_1} = x_4$ should no longer be a basic variable. Its place as basic variable is taken by $x_3$, so that $\beta = (3, 5)$ and $\nu = (1, 2, 4)$.

The corresponding basic matrix is $\mathbf{A}_\beta = \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$, while $\mathbf{A}_\nu = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

The values of the current basic variables are given by $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where the vector $\bar{\mathbf{b}}$ is calculated from the system $\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e.

$$\begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

The vector $\mathbf{y}$ with simplex multipliers is obtained from the system $\mathbf{A}_\beta^\mathsf{T} \mathbf{y} = \mathbf{c}_\beta$, i.e.

$$\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ with the solution } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Then the reduced cost for the non-basic variable is obtained from

$$\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T} \mathbf{A}_\nu = (0,\ 0,\ 1) - (-2,\ 1) \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = (3,\ 1,\ 3).$$

Since $\mathbf{r}_\nu \geq \mathbf{0}$ the current feasible basic solution is optimal.

Thus, the BFS defined by $\mathbf{x}_\beta = \bar{\mathbf{b}}$ and $\mathbf{x}_\nu = \mathbf{0}$ is an optimal solution, which means that $x_1 = 0,\ x_2 = 0,\ x_3 = 2,\ v_1 = 0,\ v_2 = 1$ is an optimal solution to the problem P3, with optimal value $v_1 + v_2 = 0 + 1 = 1$.

We shall now motivate that since the optimal value of P3 is $> 0$, there is no feasible solution to P2.

Let $\mathbf{A} = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 4 \end{bmatrix}$, $\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$,

and assume that there is at least one feasible solution to P2.
This means that there is a vector $\tilde{\mathbf{x}} \in I\!R^3$ which satisfies $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{b}$ and $\tilde{\mathbf{x}} \geq \mathbf{0}$.

Let $\tilde{\mathbf{v}} = \mathbf{0} \in I\!R^2$. Then $\begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{v}} \end{pmatrix} \in I\!R^5$ is a feasible solution to P3, since

$$[\,\mathbf{A} \quad \mathbf{I}\,] \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{v}} \end{pmatrix} = \mathbf{A}\tilde{\mathbf{x}} + \mathbf{I}\,\tilde{\mathbf{v}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{b} \text{ and } \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{v}} \end{pmatrix} \geq \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

Further, this feasible solution to P3 has the objective value $\tilde{v}_1 + \tilde{v}_2 = 0 + 0 = 0$.

But this is a contradiction, since according to above the optimal value of P3 is $= 1 > 0$.

The conclusion is that there is no feasible solution to P2.

**3.(a)** The constraints (1) kan be written on the form $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{x} = \begin{pmatrix} x_{sw} \\ x_{nw} \\ x_{ne} \\ x_{se} \end{pmatrix}, \quad \mathbf{A} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_s \\ b_w \\ b_n \\ b_e \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2b_s \\ 2b_w \\ 2b_n \\ 2b_e \end{pmatrix}.$$

We use elementary row operations to transform $\mathbf{Ax} = \mathbf{b}$ to reduced row echelon form:

$$\frac{1}{2}\left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 2b_s \\ 1 & 1 & 0 & 0 & 2b_w \\ 0 & 1 & 1 & 0 & 2b_n \\ 0 & 0 & 1 & 1 & 2b_e \end{array}\right] \longrightarrow \cdots \longrightarrow \frac{1}{2}\left[\begin{array}{cccc|l} 1 & 0 & 0 & 1 & 2b_s \\ 0 & 1 & 0 & -1 & 2b_w - 2b_s \\ 0 & 0 & 1 & 1 & 2b_n - 2b_w + 2b_s \\ 0 & 0 & 0 & 0 & 2b_e - 2b_n + 2b_w - 2b_s \end{array}\right]$$

From this reduced row echelon form, it follows that:

If $2b_e - 2b_n + 2b_w - 2b_s \neq 0$ then there is no solution to $\mathbf{Ax} = \mathbf{b}$.

If $2b_e - 2b_n + 2b_w - 2b_s = 0$ then the general solution to $\mathbf{Ax} = \mathbf{b}$ is obtained by letting $x_{se} = v$ (an arbitrary number), whereafter $x_{sw} = 2b_s - v$, $x_{nw} = 2b_w - 2b_s + v$, and $x_{ne} = 2b_n - 2b_w + 2b_s - v$, which can be written

$$\mathbf{x} = \begin{pmatrix} x_{sw} \\ x_{nw} \\ x_{ne} \\ x_{se} \end{pmatrix} = \begin{pmatrix} 2b_s \\ 2b_w - 2b_s \\ 2b_n - 2b_w + 2b_s \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} v = \bar{\mathbf{x}} + \mathbf{z}\, v,$$

for $v \in I\!R$, where $\bar{\mathbf{x}}$ is one solution to $\mathbf{Ax} = \mathbf{b}$, and $\mathbf{z}$ is a basis for the null-space of $\mathbf{A}$.

Since different values of $v$ give different solutions $\mathbf{x}$, there is an infinite number of solutions to $\mathbf{Ax} = \mathbf{b}$ if $2b_e - 2b_n + 2b_w - 2b_s = 0$. But $2b_e - 2b_n + 2b_w - 2b_s = 0 \Leftrightarrow \mathbf{w}^\mathsf{T}\mathbf{b} = 0$.

**3.(b)** Now we should solve the problem: minimize $\frac{1}{2}(\mathbf{x} - \tilde{\mathbf{b}})^\mathsf{T}(\mathbf{x} - \tilde{\mathbf{b}})$ subject to $\mathbf{Ax} = \mathbf{b}$,

$$\text{where} \quad \mathbf{b} = \begin{pmatrix} b_s \\ b_w \\ b_n \\ b_e \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 2 \end{pmatrix} \quad \text{and} \quad \tilde{\mathbf{b}} = \begin{pmatrix} b_{sw} \\ b_{nw} \\ b_{ne} \\ b_{se} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} b_s + b_w \\ b_n + b_w \\ b_n + b_e \\ b_s + b_e \end{pmatrix} = \begin{pmatrix} 2 \\ 3.5 \\ 3 \\ 1.5 \end{pmatrix}.$$

Since $\frac{1}{2}(\mathbf{x} - \tilde{\mathbf{b}})^\mathsf{T}(\mathbf{x} - \tilde{\mathbf{b}}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{x} - \tilde{\mathbf{b}}^\mathsf{T}\mathbf{x} + \frac{1}{2}\tilde{\mathbf{b}}^\mathsf{T}\tilde{\mathbf{b}}$, this is a problem of the form

$$\text{minimize } \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} + c_0 \ \text{ subject to } \ \mathbf{Ax} = \mathbf{b},$$

with $\mathbf{A}$ and $\mathbf{b}$ as above, $\mathbf{c} = -\tilde{\mathbf{b}}$, and $\mathbf{H} = \mathbf{I}$ = the $4{\times}4$ identity matrix. ($c_0$ can be ignored.)

Since a nullspace method should be used, we use the above transformation $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{z}\, v$, where $\bar{\mathbf{x}} = (2, 4, 4, 0)^\mathsf{T}$ and $\mathbf{z} = (-1,\ 1, -1,\ 1)^\mathsf{T}$.

Changing variables from $\mathbf{x} \in I\!R^4$ to $v \in I\!R$ leads to a quadratic objective function which is uniquely minimized by the solution $\hat{v}$ to the system $(\mathbf{z}^\mathsf{T}\mathbf{H}\,\mathbf{z})\, v = -\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, provided that $\mathbf{z}^\mathsf{T}\mathbf{H}\,\mathbf{z}$ is positive definite ($> 0$ in this one-variable case).

We get that $\mathbf{z}^\mathsf{T}\mathbf{H}\,\mathbf{z} = \mathbf{z}^\mathsf{T}\mathbf{z} = 4 > 0$ and $-\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}) = -\mathbf{z}^\mathsf{T}(\bar{\mathbf{x}} - \tilde{\mathbf{b}}) = 2$,
so the unique solution to $(\mathbf{z}^\mathsf{T}\mathbf{H}\,\mathbf{z})\, v = -\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ is $\hat{v} = 0.5$,
and thus the unique optimal solution to the considered QP problem is

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = (2, 4, 4, 0)^\mathsf{T} + 0.5\,(-1,\ 1, -1,\ 1)^\mathsf{T} = (1.5,\ 4.5,\ 3.5,\ 0.5)^\mathsf{T}$$

**4.(a)** Let $c_1 = 9$, $c_2 = 16$ and $c_3 = 25$.

Then the objective function can be written $f(\mathbf{x}) = \sum_{j=1}^{3} ( \frac{1}{x_j} + c_j x_j )$,

while the constraint function for the explicit constraint is $g(\mathbf{x}) = -b + \sum_{j=1}^{3} \frac{1}{x_j}$.

The Lagrange function becomes $L(\mathbf{x}, y) = f(\mathbf{x}) + yg(\mathbf{x}) = -by + \sum_{j=1}^{3} ( c_j x_j + \frac{1+y}{x_j} )$.

The Lagrange relaxed problem $\text{PR}_y$ is defined, for a given $y \geq 0$, as follows:

$$\text{PR}_y : \text{ minimize } L(\mathbf{x}, y) \text{ subject to } \mathbf{x} \in X.$$

But this problem separates into one problem for each variable $x_j$:

$$\text{minimize } \ell_j(x_j) = c_j x_j + \frac{1+y}{x_j} \text{ subject to } x_j > 0.$$

We have that $\ell_j'(x_j) = c_j - \frac{1+y}{x_j^2}$ and $\ell_j''(x_j) = \frac{2(1+y)}{x_j^3} > 0$ for all $x_j > 0$, since $y \geq 0$.
Thus, $\ell_j(x_j)$ is strictly convex. Further, the equation $\ell_j'(x_j) = 0$ has the unique solution

$$\tilde{x}_j(y) = \frac{\sqrt{1+y}}{\sqrt{c_j}}, \text{ which is} > 0, \text{ since } y \geq 0.$$

Therefore, the optimal solution to the Lagrange relaxed problem $\text{PR}_y$ is

$$\tilde{\mathbf{x}}(y) = (\tilde{x}_1(y), \tilde{x}_2(y), \tilde{x}_3(y))^{\mathsf{T}} = \left( \frac{\sqrt{1+y}}{3}, \frac{\sqrt{1+y}}{4}, \frac{\sqrt{1+y}}{5} \right)^{\mathsf{T}}.$$

The dual objective function is then given by

$$\varphi(y) = L(\tilde{\mathbf{x}}(y), y) = -by + \sum_{j=1}^{n} \ell_j(\tilde{x}_j(y)) = \cdots = -by + 24\sqrt{1+y},$$

with $\varphi'(y) = -b + \frac{12}{\sqrt{1+y}}$ and $\varphi''(y) = \frac{-6}{(1+y)^{3/2}} < 0$ for all $y \geq 0$.

**4.(b)**. Assume that $b = 6$.

Then the unique solution to $\varphi'(y) = 0$ is $\hat{y} = 3$.
This implies, since $\varphi$ is strictly concave for $y \geq 0$, that $\varphi(\hat{y}) = \varphi(3) > \varphi(y)$ for all $y \geq 0$.

The corresponding primal solution is $\hat{\mathbf{x}} = \tilde{\mathbf{x}}(\hat{y}) = (\tilde{x}_1(3), \tilde{x}_2(3), \tilde{x}_3(3))^{\mathsf{T}} = (2/3,\, 2/4,\, 2/5)^{\mathsf{T}}$.

It should now be verified that $\hat{\mathbf{x}}$ and $\hat{y}$ satisfy the global optimality conditions (GOC):

GOC-1 is satisfied since $\hat{\mathbf{x}} = \tilde{\mathbf{x}}(\hat{y})$.
GOC-2 is satisfied since $g(\hat{\mathbf{x}}) = -6 + 3/2 + 4/2 + 5/2 = 0 \leq 0$.
GOC-3 is satisfied since $\hat{y} = 3 \geq 0$.
GOC-4 is satisfied since $\hat{y} g(\hat{\mathbf{x}}) = 3 \cdot 0 = 0$.

As an additional verification of optimality, we check if $f(\hat{\mathbf{x}}) = \varphi(\hat{y})$ :

$f(\hat{\mathbf{x}}) = 3/2 + 4/2 + 5/2 + 6 + 8 + 10 = 30$ and $\varphi(\hat{y}) = -6 \cdot 3 + 24\sqrt{1+3} = 30$.

**4.(c).** Assume that $b = 18$.

Then $\varphi'(y) = -18 + \dfrac{12}{\sqrt{1+y}} < 0$ for all $y \geq 0$,

which implies that $\varphi$ is strictly decreasing for all $y \geq 0$,
which in turn implies that $\varphi(0) > \varphi(y)$ for all $y \geq 0$.

Thus, when $b = 18$, the optimal dual solution is $\hat{y} = 0$.

The corresponding primal solution is $\hat{\mathbf{x}} = \tilde{\mathbf{x}}(\hat{y}) = (\tilde{x}_1(0),\, \tilde{x}_2(0),\, \tilde{x}_3(0))^\mathsf{T} = (1/3,\, 1/4,\, 1/5)^\mathsf{T}$.

It should now be verified that $\hat{\mathbf{x}}$ and $\hat{y}$ satisfy the global optimality conditions (GOC):

GOC-1 is satisfied since $\hat{\mathbf{x}} = \tilde{\mathbf{x}}(\hat{y})$.
GOC-2 is satisfied since $g(\hat{\mathbf{x}}) = -18 + 3 + 4 + 5 = -6 \leq 0$.
GOC-3 is satisfied since $\hat{y} = 0 \geq 0$.
GOC-4 is satisfied since $\hat{y}g(\hat{\mathbf{x}}) = 0 \cdot (-6) = 0$.

As an additional verification of optimality, we check if $f(\hat{\mathbf{x}}) = \varphi(\hat{y})$ :

$f(\hat{\mathbf{x}}) = 3 + 4 + 5 + 3 + 4 + 5 = 24$ and $\varphi(\hat{y}) = -10 \cdot 0 + 24\sqrt{1+0} = 24$.

**5.** The objective function is $f(\mathbf{x}) = x_1^2 x_2^2 + 2x_1^2 + 2x_2^2 - 12x_1 - 12x_2$.

The gradient (as column vector) $\nabla f(\mathbf{x})^\mathsf{T}$ and the Hessian matrix $\mathbf{F}(\mathbf{x})$ of $f$ become

$$\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} 2x_1x_2^2 + 4x_1 - 12 \\ 2x_1^2x_2 + 4x_2 - 12 \end{pmatrix} \text{ and } \mathbf{F}(\mathbf{x}) = \begin{bmatrix} 2x_2^2 + 4 & 4x_1x_2 \\ 4x_1x_2 & 2x_1^2 + 4 \end{bmatrix}.$$

**5.(a).** The starting point for Newton's method is given by

$\mathbf{x}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Then $f(\mathbf{x}^{(1)}) = -10$, $\nabla f(\mathbf{x}^{(1)})^\mathsf{T} = \begin{pmatrix} -8 \\ -12 \end{pmatrix}$ and $\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$.

$\mathbf{F}(\mathbf{x}^{(1)})$ is positive definite since $4 > 0$, $6 > 0$ and $4 \cdot 6 - 0 \cdot 0 > 0$.
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. $\begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix} \mathbf{d} = \begin{pmatrix} 8 \\ 12 \end{pmatrix}$, with the solution $\mathbf{d}^{(1)} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

First, try $t_1 = 1$, so that $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

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

Next, 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 $f(\mathbf{x}^{(2)}) = -22 < f(\mathbf{x}^{(1)})$, so $t_1 = 0.5$ is accepted, and the first iteration is completed.

**5.(b).** Since each of the sets $C_i$, $i = 1, 2, 3$, has interior points (e.g. $\mathbf{x} = \mathbf{0}$), $f$ is convex

on $C_i$ if and only if $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 2x_2^2 + 4 & 4x_1x_2 \\ 4x_1x_2 & 2x_1^2 + 4 \end{bmatrix}$ is positive semidefinite for all $x \in C_i$.

The diagonal elements in $\mathbf{F}(\mathbf{x})$ are always $> 0$, so $\mathbf{F}(\mathbf{x})$ is positive semidefinite for all $x \in C_i$ if and only if $(2x_2^2 + 4)(2x_1^2 + 4) - (4x_1x_2)(4x_1x_2) \geq 0$ for all $\mathbf{x} \in C_i$.

In particular, if $x_1 = x_2$ then $(2x_2^2 + 4)(2x_1^2 + 4) - (4x_1x_2)(4x_1x_2) = (2x_1^2 + 4)^2 - (4x_1^2)^2$, which is $< 0$ if $4 < 2x_1^2$. Thus, if $\mathbf{x} = (x_1, x_2)^\mathsf{T} = (1.5,\, 1.5)^\mathsf{T}$ which satisfies $\mathbf{x}^\mathsf{T}\mathbf{x} = 4.5$, then $\mathbf{F}(\mathbf{x})$ is not positive semidefinite. This shows that $f$ is not convex on $C_3$.

Assume that $\mathbf{x} \in C_2$. Then $4 > \mathbf{x}^\mathsf{T}\mathbf{x} = x_1^2 + x_2^2$, and thus

$(2x_2^2 + 4)(2x_1^2 + 4) - (4x_1x_2)(4x_1x_2) > (3x_2^2 + x_1^2)(3x_1^2 + x_2^2) - (4x_1x_2)(4x_1x_2) =$

$3x_1^4 + 3x_2^4 + 10x_1^2x_2^2 - 16x_1^2x_2^2 = 3\,(x_1^4 + x_2^4 - 2x_1^2x_2^2) = 3\,(x_1^2 - x_2^2)^2 \geq 0$.

This shows that $f$ is convex on $C_2$.

Finally, since $C_1$ is a convex subset of $C_2$, $f$ is convex on $C_1$.

**5.(c)** Assume that $\tilde{\mathbf{x}}$ is a local optimum solution and $\tilde{x}_1 \neq \tilde{x}_2$. Then

$$\nabla f(\tilde{\mathbf{x}})^{\mathsf{T}} = \begin{pmatrix} 2\tilde{x}_1\tilde{x}_2^2 + 4\tilde{x}_1 - 12 \\ 2\tilde{x}_1^2\tilde{x}_2 + 4\tilde{x}_2 - 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

By subtracting the first equations from the second, we obtain that $2\tilde{x}_1\tilde{x}_2(\tilde{x}_1 - \tilde{x}_2) + 4(\tilde{x}_2 - \tilde{x}_1) = 0$, i.e. that $(2\tilde{x}_1\tilde{x}_2 - 4)(\tilde{x}_1 - \tilde{x}_2) = 0$.

This implies, since $\tilde{x}_1 \neq \tilde{x}_2$, that $2\tilde{x}_1\tilde{x}_2 = 4$.
If this is plugged into the above two equations, we obtain that

$$\nabla f(\tilde{\mathbf{x}})^{\mathsf{T}} = \begin{pmatrix} 4\tilde{x}_2 + 4\tilde{x}_1 - 12 \\ 4\tilde{x}_1 + 4\tilde{x}_2 - 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ i.e. that } \tilde{x}_1 + \tilde{x}_2 = 3,$$

which together with $2\tilde{x}_1\tilde{x}_2 = 4$ implies that $(\tilde{x}_1, \tilde{x}_2)^{\mathsf{T}} = (2,1)^{\mathsf{T}}$ or $(1,2)^{\mathsf{T}}$.

On the other hand, if $(\tilde{x}_1, \tilde{x}_2)^{\mathsf{T}} = (2,1)^{\mathsf{T}}$ or $(1,2)^{\mathsf{T}}$ then $\nabla f(\tilde{\mathbf{x}})^{\mathsf{T}} = \mathbf{0}$ and

$$\mathbf{F}(\tilde{\mathbf{x}}) = \begin{bmatrix} 6 & 8 \\ 8 & 12 \end{bmatrix} \text{ or } \begin{bmatrix} 12 & 8 \\ 8 & 6 \end{bmatrix}, \text{ which are both positive definite.}$$

Thus, the two points $(2,1)^{\mathsf{T}}$ and $(1,2)^{\mathsf{T}}$ are local optimal solutions.

**5.(d)** Assume that $\tilde{\mathbf{x}}$ is a local optimum solution and $\tilde{x}_1 = \tilde{x}_2 = t$.
Then two things must hold:

First, $\nabla f(\tilde{\mathbf{x}})^{\mathsf{T}} = \begin{pmatrix} 2t^3+4t-12 \\ 2t^3+4t-12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, i.e. $2t^3+4t-12 = 0$.

Second, $\mathbf{F}(\tilde{\mathbf{x}}) = \begin{bmatrix} 2\,t^2+4 & 4\,t^2 \\ 4\,t^2 & 2\,t^2+4 \end{bmatrix}$ is positive semidefinite, i.e. $4 \geq 2\,t^2$.

Since $p(t) = 2t^3+4t - 12$ is a strictly increasing polynomial with $p(1) < 0$ and $p(2) > 0$, there is a unique solution to the equation $p(t) = 0$, and this solution satisfies $1 < t < 2$.
Then the above requirement that $4 \geq 2\,t^2$ implies that $1 < t \leq \sqrt{2}$.
But since both $p(1) < 0$ and $p(\sqrt{2}) = 8(\sqrt{2} - 1.5) < 0$, there is no solution to $p(t) = 0$ with $t \in (1, \sqrt{2}\,]$.

Thus, the assumption that $\tilde{\mathbf{x}}$ is a local optimum solution and $\tilde{x}_1 = \tilde{x}_2$ leads to a contradiction. Therefore, there is no such $\tilde{\mathbf{x}}$.