Solutions to exam in SF1811 Optimization, March 2016

1.(a)

We note that each column in the matrix $\mathbf{A}$ contains one "+1" and one "-1", while all the other elements in the column are zeros. We also note that the sum of the elements in the vector $\mathbf{b}$ is zero. These observations imply that the LP problem is in fact a balanced network flow problem with 5 nodes (one for each row in $\mathbf{A}$) and 6 directed arcs (one for each column in $\mathbf{A}$). The network corresponding to the given $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{c}$ in this exercise can be illustrated by FIGURE 1 below, where the supply at the nodes (i.e. the components in the vector $\mathbf{b}$), and the unit costs of the arcs (i.e. the components in the vector $\mathbf{c}$) are written in the figure. Arcs from Node1 are directed from left to right, while arcs from Node2 are directed from right to left. Negative supply means demand. If the flow in the arc from node i to node j is denoted x_{ij}, the variable vector is $\mathbf{x} = (x_{13}, x_{14}, x_{15}, x_{23}, x_{24}, x_{25})^\mathsf{T}$.

1.(b)

The suggested solution $\tilde{\mathbf{x}} = (400, 300, 0, 0, 200, 600)^\mathsf{T}$ can be illustrated by the spanning tree in FIGURE 2 below, with the arc-flows written on the arcs.

```
            Node3
            -400
              O                               O
             / \                             /
            /   \                           /
           /     \                         /
          5       q                      400
         /         \                     /
        /           \                   /
Node1  /    Node4    \  Node2          /
 +700 O---3---O---2---O +800          O--300--O--200--O
       \    -500     /                               /
        \           /                               /
         \         /                               /
          7       4                              600
           \     /                               /
            \   /                               /
             \ /                               /
              O                               O
            Node5
            -600

           FIGURE 1                        FIGURE 2
```

The suggested solution is clearly feasible, since there is flow balance in each of the four nodes and all arc-flows are non-negative. It remains to show that it is optimal.

The simplex multipliers y_i for the five nodes are calculated by $y_5 = 0$ and $y_i - y_j = c_{ij}$ for all arcs (i, j) in the spanning tree. By using FIGURE 3 below, the y_i are calculated in the order $y_5 = 0$, $y_2 = 4$, $y_4 = 2$, $y_1 = 5$ and $y_3 = 0$.

Then the reduced cost for the two non-basic variables are calculated by $r_{ij} = c_{ij} - y_i + y_j$, see FIGURE 4 below.

```
              O y3=0                         y3=0 O
             /                                     \
            /                                       \
           /                                         \
          5                                      r23=q-4+0=q-4
         /                                             \
        /                                               \
       /    y4=2                                         \
y1=5 O---3---O---2---O y2=4                   O y1=5      y2=4 O
                    /                          \
                   /                            \
                  /                              \
                 4                           r15=7-5+0=2
                /                                  \
               /                                    \
              /                                      \
             O y5=0                                   O y5=0

          FIGURE 3                              FIGURE 4
```

If $q = 5$ then both $r_{15} \geq 0$ and $r_{23} \geq 0$, so that the suggested solution $\tilde{\mathbf{x}}$ is optimal.

1.(c)

If $q = 3$, then the reduced cost r_{23} in FIGURE 4 becomes $r_{23} = q - 4 = -1 < 0$, which means that we should let the currently non-basic variable x_{23} increase from zero. Thus, let $x_{23} = t$, where t increase from zero. The current basic variables dependence of t is illustrated in FIGURE 5 below. It is clear that t may increase to at most 200. Then the new basic feasible solution $\hat{\mathbf{x}}$ in FIGURE 6 is obtained.

```
              0                        0
             / \                      / \
            /   \                    /   \
           /     \                  /     \
        400-t   +t               200      200
         /        \              /          \
        /          \            /            \
       /            \          /              \
      0-------0-------0       0--500--0        0
        300+t  200-t /                        /
                    /                        /
                   /                        /
                600                      600
                 /                        /
                /                        /
               /                        /
              0                        0

           FIGURE 5                 FIGURE 6
```

Again, the simplex multipliers y_i for the nodes are calculated by $y_5 = 0$ and $y_i - y_j = c_{ij}$ for all arcs (i, j) in the spanning tree. By using FIGURE 7 below, the y_i are calculated in the order $y_5 = 0$, $y_2 = 4$, $y_3 = 4 - q$, $y_1 = 9 - q$ and $y_4 = 6 - q$.
Then the reduced cost for the two non-basic variables are calculated by $r_{ij} = c_{ij} - y_i + y_j$, see FIGURE 8 below.

```
               0 y3=4-q
              / \
             /   \
            /     \
           5       q
          /         \
         /           \
        /    y4=6-q   \            y1=9-q  y4=6-q  y2=4
y1=9-q 0---3---0       0 y2=4        0       0-------0
                      /               \      r24=2-4+(6-q)=4-q
                     /                 \
                    /                   \
                   4               r15=7-(9-q)+0=q-2
                  /                       \
                 /                         \
                /                           \
               0 y5=0                        0 y5=0

            FIGURE 7                      FIGURE 8
```

Since $q = 3$, both $r_{15} \geq 0$ and $r_{24} \geq 0$, so that $\hat{\mathbf{x}}$ in FIGURE 6 is optimal.

1.(d)

It follows from FIGURE 4 that the solution $\tilde{\mathbf{x}}$ in FIGURE 2 is optimal whenever $q \geq 4$. Therefore, the optimal value of the considered LP problem is in this case

$\mathbf{c}^{\mathsf{T}}\tilde{\mathbf{x}} = 5 \cdot 400 + 3 \cdot 300 + 2 \cdot 200 + 4 \cdot 600 = 5700.$ (If $q \geq 4$.)

It follows from FIGURE 8 that the solution $\hat{\mathbf{x}}$ in FIGURE 6 is optimal whenever $2 \leq q \leq 4$. Therefore, the optimal value of the considered LP problem is in this case

$\mathbf{c}^{\mathsf{T}}\hat{\mathbf{x}} = 5 \cdot 200 + 3 \cdot 500 + q \cdot 200 + 4 \cdot 600 = 4900 + 200q.$ (If $2 \leq q \leq 4$.)

The required graph of the optimal value (vertical axis) for different values on q (horizontal axis) thus looks as follows:

```
| 5700              ---------
|                  /
|                 /
|                /
|               /
|              /
|             /
|            /
| 5300      /
|
|
|           2   3   4   5   6
|------------------------------
```

2.(a) We have a QP problem with equality constraints, i.e. a problem of the form minimize $\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$ subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$,

with $\mathbf{H} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{A} = [\,1 \quad 2 \quad 3\,]$ and $\mathbf{b} = 4$.

The general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, i.e. $x_1 + 2x_2 + 3x_3 = 4$, is given by

$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \cdot v_1 + \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \cdot v_2$, for arbitrary values on v_1 and v_2,

which means that $\bar{\mathbf{x}} = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}$ is a feasible solution, and $\mathbf{Z} = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ is a matrix whos columns form a basis for the nullspace of $\mathbf{A}$.

After the variable change $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$ we should solve the system $(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, provided that $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z}$ is at least positive semidefinite.

We have that $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z} = \begin{bmatrix} 4 & 4 \\ 4 & 6 \end{bmatrix}$, which is positive definite since $4 > 0$, $6 > 0$, $4 \cdot 6 - 4 \cdot 4 > 0$.

The system $(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ becomes $\begin{bmatrix} 4 & 4 \\ 4 & 6 \end{bmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, with the (unique) solution $\hat{\mathbf{v}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Thus, $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ is the (unique) optimal solution.

2.(b) Now the constraint matrix $\mathbf{A}$ should be changed to $\mathbf{A} = [\,1 \quad -2 \quad 3\,]$, which implies that the nullspace matrix $\mathbf{Z}$ is changed to $\mathbf{Z} = \begin{bmatrix} 2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$, while the vector $\bar{\mathbf{x}}$ is unchanged.

Then $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z} = \begin{bmatrix} -4 & 0 \\ 0 & 6 \end{bmatrix}$, which is not positive semidefinite.

Thus, there is no optimal solution to the problem in this case.

With the variable change $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}$, we get (using that $\mathbf{c} = \mathbf{0}$) that

$f(\mathbf{x}) = f(\bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}) = \frac{1}{2}(\bar{\mathbf{x}} + \mathbf{Z}\mathbf{v})^\mathsf{T}\mathbf{H}(\bar{\mathbf{x}} + \mathbf{Z}\mathbf{v}) = \frac{1}{2}\bar{\mathbf{x}}^\mathsf{T}\mathbf{H}\bar{\mathbf{x}} + \bar{\mathbf{x}}^\mathsf{T}\mathbf{H}\mathbf{Z}\mathbf{v} + \frac{1}{2}\mathbf{v}^\mathsf{T}(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} =$

$= 0 - 4v_1 - 4v_2 - 2v_1^2 + 3v_2^2$.

So by letting $v_1 = t$ and $v_2 = 0$, the objective value $\to -\infty$ when $t \to \infty$.

This corresponds to letting $\mathbf{x}(t) = \bar{\mathbf{x}} + t \cdot \mathbf{d}$, where $\mathbf{d} = (2, 1, 0)^\mathsf{T}$ = the first column in $\mathbf{Z}$.

Then $\mathbf{A}\mathbf{x}(t) = \mathbf{b}$ for all $t \in \mathbb{R}$ and $f(\mathbf{x}(t)) = -4t - 2t^2 \to -\infty$ when $t \to \infty$.

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

$$\begin{aligned} \text{minimera} \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}\,, \\ & \mathbf{x} \geq \mathbf{0}\,, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 & 1 & 0 \\ 2 & 2 & 1 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 180 \\ 120 \end{pmatrix}$ and $\mathbf{c}^\mathsf{T} = (-3,\, -4,\, -2,\, 0,\, 0)$.

The starting solution should have the basic variables 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 $\mathbf{A}_\nu = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 1 \end{bmatrix}$.

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.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 180 \\ 120 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 180 \\ 120 \end{pmatrix}.$$

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

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

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

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

Since $r_{\nu_2} = r_2 = -4$ is smallest, and < 0, we let x_2 become the new basic variable.

Then we should calculate the vector $\bar{\mathbf{a}}_2$ from the system $\mathbf{A}_\beta \, \bar{\mathbf{a}}_2 = \mathbf{a}_2$, i.e.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{a}}_2 = \begin{pmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

The largest permitted value of the new basic variable x_2 is then given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i2}} \mid \bar{a}_{i2} > 0 \right\} = \min \left\{ \frac{180}{2},\, \frac{120}{2} \right\} = \frac{120}{2} = \frac{\bar{b}_2}{\bar{a}_{12}}.$$

Minimizing index is $i = 2$, which implies that $x_{\beta_2} = x_5$ should no longer be a basic variable. Its place as basic variable is taken by x_2, so that $\beta = (4,2)$ and $\nu = (1,5,3)$.

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

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.

$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 180 \\ 120 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 60 \\ 60 \end{pmatrix}.$$

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

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

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

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

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

Thus, $\mathbf{x} = (0, 60, 0, 60, 0)^\mathsf{T}$ is an optimal solution, with optimal value $\mathbf{c}^\mathsf{T}\mathbf{x} = -240$.

3.(b) If the primal problem is on the standard form

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

the corresponding dual problem is: maximize $\mathbf{b}^\mathsf{T}\mathbf{y}$ subject to $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$, which becomes

$$\begin{array}{rrcrcr} \text{maximize} & 180y_1 & + & 120y_2 & & \\ \text{subject to} & y_1 & + & 2y_2 & \leq & -3, \\ & 2y_1 & + & 2y_2 & \leq & -4, \\ & 2y_1 & + & y_2 & \leq & -2, \\ & y_1 & & & \leq & 0, \\ & & & y_2 & \leq & 0. \end{array}$$

This dual problem can be illustrated by drawing the constraints and some level lines (orthogonal to the vector $(180, 120)^\mathsf{T}$) for the objective function, in a coordinate system with y_1 and y_2 on the axes. (The figure is omitted here.)

From this figure it is seen that the point $\mathbf{y} = (0,\ -2)^\mathsf{T}$ is an optimal dual solution.

This is consistent with the well-known fact that an optimal dual solution is given by the vector $\mathbf{y}$ of "simplex multipliers" for the optimal basic solution in (a) above.

One more optimality check: $\mathbf{y} = (0,\ -2)^\mathsf{T}$ satisfies the dual constraints, with dual objective value $180y_1 + 120y_2 = -240 =$ the optimal value of the primal problem!

3.(c)

The complementarity conditions become:

$$
\begin{aligned}
&x_1 + 2x_2 + 2x_3 + x_4 = 180,\\
&2x_1 + 2x_2 + \ x_3 + x_5 = 120,\\
&x_j \geq 0,\ j = 1, \dots, 5,\\
&y_1 + 2y_2 + 3 \leq 0,\\
&2y_1 + 2y_2 + 4 \leq 0,\\
&2y_1 + \ y_2 + 2 \leq 0,\\
&y_1 \leq 0,\\
&y_2 \leq 0,\\
&x_1\,(\ y_1 + 2y_2 + 3) = 0,\\
&x_2\,(2y_1 + 2y_2 + 4) = 0,\\
&x_3\,(2y_1 + \ y_2 + 2) = 0,\\
&x_4\,y_1 = 0,\\
&x_5\,y_2 = 0.
\end{aligned}
$$

From (b) above, we know that $y_1 = 0$, $y_2 = -2$ is an optimal dual solution.
If these values are plugged into the complementarity conditions above, it follows that $\mathbf{x}$ is an optimal primal solution if and only if it satisfies the following:

$$
\begin{aligned}
&x_1 + 2x_2 + 2x_3 + x_4 = 180,\\
&2x_1 + 2x_2 + \ x_3 + x_5 = 120,\\
&x_j \geq 0,\ j = 1, \dots, 5,\\
&x_1\,(0 - 2 + 3) = 0,\\
&x_2\,(0 - 4 + 4) = 0,\\
&x_3\,(0 - 2 + 2) = 0,\\
&0x_4 = 0,\\
&-2x_5 = 0,
\end{aligned}
$$

which may be simplified to

$$
\begin{aligned}
&2x_2 + 2x_3 + x_4 = 180,\\
&2x_2 + \ x_3 \qquad\ \ = 120,\\
&x_1 = 0,\\
&x_5 = 0,\\
&x_j \geq 0,\ j = 2, 3, 4,
\end{aligned}
$$

The general solution to $\begin{array}{l} 2x_2 + 2x_3 + x_4 = 180 \\ 2x_2 + \ x_3 \qquad\ = 120 \end{array}$ is $\begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 30{+}t/2 \\ 60{-}t \\ t \end{pmatrix}$ for $t \in I\!R$,
whereafter the constraints $x_j \geq 0,\ j = 2, 3, 4$ imply that $0 \leq t \leq 60$.

Thus, the complete set of optimal solution to the primal problem is given by

$(x_1, x_2, x_3, x_4, x_5)^{\mathsf{T}} = (0,\ 30{+}t/2,\ 60{-}t,\ t,\ 0)^{\mathsf{T}}$ for $t \in [\,0,\ 60\,]$.

Two of these optimal solutions are basic feasible solutions, namely $\tilde{\mathbf{x}} = (0,\ 30,\ 60,\ 0,\ 0)^{\mathsf{T}}$ and $\hat{\mathbf{x}} = (0,\ 60,\ 0,\ 60,\ 0)^{\mathsf{T}}$, and every optimal solution is a convex combination of $\tilde{\mathbf{x}}$ and $\hat{\mathbf{x}}$.

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

Then the gradient and Hessian of f becomes:

$$\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} 3x_1^2 - 3x_2 \\ 3x_2^2 - 3x_1 \end{pmatrix} \text{ and } \mathbf{F}(\mathbf{x}) = \begin{bmatrix} 6x_1 & -3 \\ -3 & 6x_2 \end{bmatrix}.$$

We will use the well known fact that a symmetric 2×2 matrix $\mathbf{H} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$

* is positive definite if and only if $a > 0$, $c > 0$ and $ac - b^2 > 0$,
* is positive semidefinite if and only if $a \geq 0$, $c \geq 0$ and $ac - b^2 \geq 0$,
which is easily verified, e.g. by an LDLT factorization.

The starting point for Newtons method is $\mathbf{x}^{(1)} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, with $f(\mathbf{x}^{(1)}) = 4$.

$\mathbf{F}(\mathbf{x}^{(1)}) = \begin{bmatrix} 12 & -3 \\ -3 & 12 \end{bmatrix}$ is positive definite since $12 > 0$, $12 > 0$ and $12 \cdot 12 - (-3) \cdot (-3) > 0$.

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. $\begin{bmatrix} 12 & -3 \\ -3 & 12 \end{bmatrix} \mathbf{d} = \begin{pmatrix} -6 \\ -6 \end{pmatrix}$, with the solution $\mathbf{d}^{(1)} = \begin{pmatrix} -2/3 \\ -2/3 \end{pmatrix}$.

First try $t_1 = 1$, so that $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1 \mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = \begin{pmatrix} 4/3 \\ 4/3 \end{pmatrix}$.

Then $f(\mathbf{x}^{(2)}) = -16/27 < f(\mathbf{x}^{(1)})$, so $t_1 = 1$ is accepted, and the first iteration is completed.

4.(b) Any local optimal solution to the problem of minimizing $f(\mathbf{x})$ without constraints

must satisfy that $\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} 3x_1^2 - 3x_2 \\ 3x_2^2 - 3x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, i.e. $x_2 = x_1^2$ and $x_1 = x_2^2$,

which imply that $x_2 = x_2^4$, i.e. $x_2(x_2^3 - 1) = 0$, with the solutions $x_2 = 0$ or $x_2 = 1$.

If $x_2 = 0$ then $x_1 = 0$ and if $x_2 = 1$ then $x_1 = 1$.

Thus, the only solutions to $\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ are $\tilde{\mathbf{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\hat{\mathbf{x}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

$\mathbf{F}(\tilde{\mathbf{x}}) = \begin{bmatrix} 0 & -3 \\ -3 & 0 \end{bmatrix}$ is not positive semidefinite, so $\tilde{\mathbf{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is *not* a local optimal solution.

$\mathbf{F}(\hat{\mathbf{x}}) = \begin{bmatrix} 6 & -3 \\ -3 & 6 \end{bmatrix}$ is positive definite, so $\hat{\mathbf{x}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ *is* a local optimal solution.

Thus, the only local optimal solution to the problem of minimizing $f(\mathbf{x})$ without any constraints is $\hat{\mathbf{x}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, with $f(\hat{\mathbf{x}}) = -1$.

(But there is no global optimal solution. If $\mathbf{x}(t) = (-t,\, 0)^\mathsf{T}$ then $f(\mathbf{x}(t)) = -t^3 \to -\infty$ when $t \to \infty$.)

4.(c) With $h(\mathbf{x}) = x_1 + x_2 + q$, the Lagrange optimality conditions become:

$$\nabla f(\mathbf{x})^\mathsf{T} + \nabla h(\mathbf{x})^\mathsf{T} u = \begin{pmatrix} 3x_1^2 - 3x_2 + u \\ 3x_2^2 - 3x_1 + u \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad h(\mathbf{x}) = x_1 + x_2 - q = 0.$$

The first two equations imply that $0 = 3(x_1^2 - x_2^2 + x_1 - x_2) = 3(x_1 - x_2)(x_1 + x_2 + 1)$,

i.e. $x_1 = x_2$ or $x_1 + x_2 + 1 = 0$ (or both).

If $q = -1$ then $x_1 + x_2 - 1 = 0$ for all feasible solutions $\mathbf{x}$,
so then the only remaining possibility is $x_1 = x_2$.
Then $h(\mathbf{x}) = 0$ implies that $x_1 = x_2 = 0.5$, whereafter $u = 0.75$.

If $q = 1$ then $x_1 + x_2 + 1 = 0$ for all feasible solutions $\mathbf{x}$, and then *every* feasible solutions $\mathbf{x}$ satisfies the Lagrange optimality conditions together with the Lagrange multiplier $u = 3(x_2 - x_1^2) = 3(x_1 - x_2^2)$!

Concerning the question of local and/or global optimality, we can use a nullspace approach as in QP. The complete set of solutions to the constraint $x_1 + x_2 + q = 0$ is

$x_1(t) = -q - t = -(q + t)$ and $x_2(t) = t$ for $t \in \mathbb{R}$.

The objective function values for these feasible solutions are

$f(\mathbf{x}(t)) = -(q+t)^3 + t^3 + 3(q+t)\,t = -q^3 + 3(q - q^2)\,t + 3(1-q)\,t^2$.

If $q = -1$ then $f(\mathbf{x}(t)) = 1 - 6\,t + 6\,t^2$, with a strikt global minimum for $t = 0.5$, which implies that $\mathbf{x}(0.5) = (0.5,\, 0.5)^\mathsf{T}$ is a global optimal solution to P1.

If $q = 1$ then $f(\mathbf{x}(t)) = 1$ for all $t \in \mathbb{R}$, which implies that *every* feasible solutions $\mathbf{x}$ to P1 is a global optimal solution to P1.

5.(a)

Since $d_1(\mathbf{x})+d_2(\mathbf{x})+d_3(\mathbf{x}) = x_1 + x_2 + b-a_1x_1-a_2x_2$, the problem becomes:

$$\begin{array}{rl} \text{minimize} & b + (1-a_1)x_1 + (1-a_2)x_2 \\ \text{subject to} & a_1x_1+a_2x_2 \leq b, \\ & x_1 \geq 0 \text{ and } x_2 \geq 0. \end{array}$$

Since $a_1^2+a_2^2 = 1$, $a_1 > 0$ and $a_2 > 0$, it follows that $a_1 < 1$ and $a_2 < 1$.
Thus, $1-a_1 > 0$ and $1-a_2 > 0$, and since both x_1 and x_2 are required to be ≥ 0, the objective function can never be less than b for any feasible solution.
With $\mathbf{x} = (0,0)^\mathsf{T}$, which is a feasible solution since $b > 0$, the objective value is $= b$.
For any other feasible solution, which has at least one $x_j > 0$, the objective value is $> b$.
Thus, $\mathbf{x} = (0,0)^\mathsf{T}$ is the unique optimal solution to the problem.

5.(b)

Since $(d_1(\mathbf{x}))^2+(d_2(\mathbf{x}))^2+(d_3(\mathbf{x}))^2 = x_1^2+x_2^2+(b-a_1x_1-a_2x_2)^2$, the problem can be written:

$$\begin{array}{rl} \text{minimize} & \|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 \\ \text{subject to} & a_1x_1+a_2x_2 \leq b, \\ & x_1 \geq 0 \text{ and } x_2 \geq 0, \end{array}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -a_1 & -a_2 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ -b \end{pmatrix}$, so that $\mathbf{A}\mathbf{x}-\mathbf{b} = \begin{pmatrix} x_1 \\ x_2 \\ b-a_1x_1-a_2x_2 \end{pmatrix}$.

Let us first try to minimize the objective function without any constraints, to see if there is an optimal solution in the interior of the triangle. This leads to the least squares problem

$$\text{minimize } \|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 \text{ subject to } \mathbf{x} \in \mathbb{R}^2.$$

which is equivalent to the normal equations $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b}$, which in this case become

$$\begin{bmatrix} 1+a_1^2 & a_1a_2 \\ a_1a_2 & 1+a_2^2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_1b \\ a_2b \end{pmatrix}.$$

Using the hint, we get that the solution to these equations is

$$\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \frac{1}{2} \begin{bmatrix} 1+a_2^2 & -a_1a_2 \\ -a_1a_2 & 1+a_1^2 \end{bmatrix} \begin{pmatrix} a_1b \\ a_2b \end{pmatrix} = \frac{b}{2} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

But then $x_1 > 0$, $x_2 > 0$ and $a_1x_1+a_2x_2 = 0.5\,b\,(a_1^2+a_2^2) = 0.5\,b < b$,
so that $\hat{\mathbf{x}}$ satisfies all the constraints (with strict inequalities).

Thus, since $\hat{\mathbf{x}} \in T$ and $\hat{\mathbf{x}}$ is the minimizing point on the whole $\mathbb{R}^2$,
$\hat{\mathbf{x}}$ must in particular be the minimizing point on T.

5.(c)
The problem of minimizing the maximal distance can be written as the following LP problem in x_1, x_2 and z:

$$\begin{array}{rl} \text{minimize} & z \\ \text{subject to} & z - x_1 \geq 0, \\ & z - x_2 \geq 0, \\ & z + a_1x_1 + a_2x_2 \geq b, \\ & -a_1x_1 - a_2x_2 \geq -b, \\ & x_1 \geq 0, \\ & x_2 \geq 0. \end{array}$$

The objective function and the first three constraints imply that, at any optimal solution, $z = \max\{\ x_1,\ x_2,\ b - a_1x_1 - a_2x_2\ \} = \max\{\ d_1(\mathbf{x}),\ d_2(\mathbf{x}),\ d_3(\mathbf{x})\ \}$.
The remaining three constraints are equivalent to that $\mathbf{x} \in T$.

The corresponding dual LP problem in $\mathbf{y} \in I\!\!R^4$ is

$$\begin{array}{rl} \text{maximize} & by_3 - by_4 \\ \text{subject to} & y_1 + y_2 + y_3 = 1, \\ & -y_1 + a_1y_3 - a_1y_4 \leq 0, \\ & -y_2 + a_2y_3 - a_2y_4 \leq 0, \\ & y_i \geq 0,\ i = 1, 2, 3, 4. \end{array}$$

After carefully looking at a figure of T, a qualified guess is that the optimal $\mathbf{x}$ is located at the point in the triangle where all three distances are equal, i.e. at the solution to

$x_1 = x_2 = b - a_1x_1 - a_2x_2$, which is $\hat{x}_1 = \hat{x}_2 = \dfrac{b}{1 + a_1 + a_2}$, and then $\hat{z} = \dfrac{b}{1 + a_1 + a_2}$.

$(\hat{x}_1,\ \hat{x}_2,\ \hat{z})^\mathsf{T}$ is an optimal solution to the primal problem if and only if there is a vector $\hat{\mathbf{y}} \in I\!\!R^4$ such that $(\hat{x}_1,\ \hat{x}_2,\ \hat{z})^\mathsf{T}$ is feasible to the primal, $\hat{\mathbf{y}}$ is feasible to the dual, and the following complementarity conditions hold:

$$\begin{array}{l} (\hat{z} - \hat{x}_1)\,\hat{y}_1 = 0, \\ (\hat{z} - \hat{x}_2)\,\hat{y}_2 = 0, \\ (\hat{z} + a_1\hat{x}_1 + a_2\hat{x}_2 - b)\,\hat{y}_3 = 0, \\ (b - a_1\hat{x}_1 - a_2\hat{x}_2)\,\hat{y}_4 = 0, \\ (-\hat{y}_1 + a_1\hat{y}_3 - a_1\hat{y}_4)\,\hat{x}_1 = 0, \\ (-\hat{y}_2 + a_2\hat{y}_3 - a_2\hat{y}_4)\,\hat{x}_2 = 0. \end{array}$$

Plugging in the above $(\hat{x}_1,\ \hat{x}_2,\ \hat{z})^\mathsf{T}$ (which is feasible to the primal), the complementarity conditions, together with the dual equality constraint, simplifies to:

$$\begin{array}{r} \hat{y}_4 = 0, \\ -\hat{y}_1 + a_1\hat{y}_3 - a_1\hat{y}_4 = 0, \\ -\hat{y}_2 + a_2\hat{y}_3 - a_2\hat{y}_4 = 0, \\ \hat{y}_1 + \hat{y}_2 + \hat{y}_3 = 1, \end{array}$$

with the solution $\hat{\mathbf{y}} = \gamma\,(a_1,\ a_2,\ 1,\ 0)^\mathsf{T}$, where $\gamma = 1/(1 + a_1 + a_2)$.
Since this $\hat{\mathbf{y}}$ is also feasible to the dual, our suggested $(\hat{x}_1,\ \hat{x}_2,\ \hat{z})^\mathsf{T}$ is optimal to the primal (and $\hat{\mathbf{y}}$ is optimal to the dual).