Solutions to exam in SF1811 Optimization, April 10, 2017

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

If x_1 and x_2 are the basic variables, then $\beta = (1,2)$ and $\nu = (3,4)$,

with $\mathbf{A}_\beta = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{A}_\nu = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}$.

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.

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

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

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

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

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

Since $r_{\nu_1} = r_3 = -1$ is smallest, and < 0, we let x_3 increase from zero.
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.

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

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

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \min \left\{ \frac{1}{1}\,, \frac{3}{1}\,, \right\} = \frac{1}{1} = \frac{\bar{b}_1}{\bar{a}_{13}}.$$

Minimizing index is $i = 1$, which implies that $x_{\beta_1} = x_1$ should no longer be a basic variable. Its place as basic variable is taken by x_3, so that $\beta = (3,2)$ and $\nu = (1,4)$.

The corresponding basic matrix is $\mathbf{A}_\beta = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, while $\mathbf{A}_\nu = \begin{bmatrix} 1 & -3 \\ 0 & -2 \end{bmatrix}$.

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.

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

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

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

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

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

Since $\mathbf{r}_\nu^{\mathsf{T}} \geq \mathbf{0}^{\mathsf{T}}$, the current BFS is optimal: $\hat{\mathbf{x}} = (0\,,2\,,1\,,0)^{\mathsf{T}}$ with $\mathbf{c}^{\mathsf{T}}\hat{\mathbf{x}} = 7$.

1.(b)
Now we consider the same LP problem as above, but with the cost vector changed from $\mathbf{c}^\mathsf{T} = (\,2,\ 2,\ 3,\ -4\,)$ to $\mathbf{c}^\mathsf{T} = (\,2,\ 2,\ 3,\ -6\,)$.

Then the first iteration becomes the same as in 1.(a). In the second iteration, when $\beta = (3, 2)$ and $\nu = (1, 4)$, the reduced costs for the non-basic variables now become
$\mathbf{r}_\nu^\mathsf{T} = \mathbf{c}_\nu^\mathsf{T} - \mathbf{y}^\mathsf{T}\mathbf{A}_\nu = (2\,, -6) - (1\,,\,1) \begin{bmatrix} 1 & -3 \\ 0 & -2 \end{bmatrix} = (1\,, -1).$

Since $r_{\nu_2} = r_4 = -1$ is smallest, and < 0, we let x_4 increase from zero.
Then we should calculate the vector $\bar{\mathbf{a}}_4$ from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_4 = \mathbf{a}_4$, i.e.
$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \end{pmatrix}$, with the solution $\bar{\mathbf{a}}_4 = \begin{pmatrix} \bar{a}_{14} \\ \bar{a}_{24} \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.

Since $\bar{\mathbf{a}}_4 \le \mathbf{0}$, the simplex method stops here, with the conclusion that there is no optimal solution to the problem.

If $x_4 = t > 0$ and $x_1 = 0$, the current basic variables become $\mathbf{x}_\beta = \bar{\mathbf{b}} - \bar{\mathbf{a}}_4 t$, i.e. $x_2 = 2{+}t$ and $x_3 = 1{+}t$, while the objective value becomes $z = \bar{z} + r_4 t = 7{-}t$.

Thus, $\mathbf{x}(t) = (0\,, 2{+}t\,, 1{+}t\,, t)^\mathsf{T}$ satisfies $\mathbf{A}\mathbf{x}(t) = \mathbf{b}$ and $\mathbf{x}(t) \ge 0$ for all $t \ge 0$, and $z(t) = \mathbf{c}^\mathsf{T}\mathbf{x}(t) = 7{-}t \to -\infty$ when $t \to +\infty$.

1.(c)
If the primal problem is on the standard form

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

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

$$\begin{array}{rrcrcl} \text{maximize} & 4y_1 & + & 3y_2 & & \\ \text{subject to} & y_1 & & & \le & 2, \\ & y_1 & + & y_2 & \le & 2, \\ & 2y_1 & + & y_2 & \le & 3, \\ & -3y_1 & - & 2y_2 & \le & c_4. \end{array}$$

If $c_4 = -4$, as in 1.(a), the the feasible region to the dual problem becomes a triangle with corner points $(0\,,\,2)^\mathsf{T}$, $(1\,,\,1)^\mathsf{T}$ and $(2\,, -1)^\mathsf{T}$. (Figure is omitted here.)

If $c_4 = -6$, as in 1.(b), then the feasible region to the dual problem becomes empty, which means that there are no feasible solutions to the dual problem.
This is consistent with the fact that the primal problem in 1.(b) has a half line $x(t)$ of feasible solutions for which $\mathbf{c}^\mathsf{T}\mathbf{x}(t) \to -\infty$. (Figure is omitted here.)

2.(a) The considered network is illustrated by FIGURE 1 below, where the supply at the nodes are written in the figure. Negative supply means demand.
The arc från Node2 to Node3 is directed from left to right.
All other arc are directed downwards in the figure.
The cost per unit flow is equal to 1 for all arcs.
The matrix **A** and the vector **b** are as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} u \\ 10-u \\ -v \\ -10+v \end{pmatrix}.$$

The equation corresponding to node 4 can be removed since it is a linear combination of the other three equations, but that is not necessary and has not been done here.

2.(b) The basic solution $\mathbf{x}$ corresponding to the spanning tree T_1 in FIGURE 2 below has been calculated as follows:
$x_{12} = u$, due to the flow balance requirement in node 1,
$x_{23} = 10$, due to the flow balance requirement in node 2,
$x_{34} = 10-v$, due to the flow balance requirement in node 3.
Then the flow balance requirement in node 4 is also fulfilled,
The basic solution corresponding to T_1 is thus $\mathbf{x} = (u,\, 0,\, 10,\, 0,\, 10-v)^\mathsf{T}$,
which is a BFS since $\mathbf{x} \geq \mathbf{0}$, and $\mathbf{c}^\mathsf{T}\mathbf{x} = 20+u-v$.

```
            Node1                           Node1
              u                               u
              o                               o
             / \                             /
            /   \                           /
           /     \                         u
          /       \                       /
Node2    /         \   Node3   Node2     /         Node3
10-u    o---------->o  -v      10-u    o--- 10 --->o -v
         \         /                               /
          \       /                               /
           \     /                            10-v
            \   /                              /
             \ /                              /
              o                              o
           -(10-v)                        -(10-v)
            Node4                          Node4
           FIGURE 1                       FIGURE 2
```

The simplex multipliers y_i for the nodes are calculated by $y_4 = 0$ and $y_i - y_j = c_{ij}$ for all arcs (i, j) in the spanning tree. Using that $c_{ij} = 1$ for all arcs, the y_i are calculated in the order $y_4 = 0,\ y_3 = y_4 + c_{34} = 1,\ y_2 = y_3 + c_{23} = 2,\ y_1 = y_2 + c_{12} = 3$.
Then the reduced cost for the two non-basic variables are calculated by $r_{ij} = c_{ij} - y_i + y_j$, i.e. $r_{13} = c_{13} - y_1 + y_3 = -1$ and $r_{24} = c_{24} - y_2 + y_4 = -1$, which implies that a lower objective value can be obtained by increasing x_{13} or x_{24}.

Let us choose to set $x_{13} = t$ (while $x_{24} = 0$) and let t increase from zero.
Then the basic variables (i.e. the arc-flows in the tree) are changed as follows:
$x_{12} = u - t$, due to the flow balance requirement in node 1,
$x_{23} = 10 - t$, due to the flow balance requirement in node 2,
$x_{34} = 10 - v$ (unchanged), due to the flow balance requirement in node 3.
Then the flow balance requirement in node 4 is also fulfilled,
By letting $t = u$ the solution $\tilde{\mathbf{x}} = (0,\, u,\, 10-u,\, 0,\, 10-v)^\mathsf{T}$ is obtained, se FIGURE 3.
This is clearly a BFS, since it corresponds to a spanning tree and $\tilde{\mathbf{x}} \geq \mathbf{0}$.
Further, $\mathbf{c}^\mathsf{T}\tilde{\mathbf{x}} = 20 - v < 20 + u - v = \mathbf{c}^\mathsf{T}\mathbf{x}$.

```
          Node1                         Node1
            u                             u
            o                             o
             \                             \
              \                             \
               u                             u
                \                             \
Node2            \  Node3     Node2            \  Node3
10-u  o-- 10-u -->o -v        10-u  o--- v-u -->o -v
                 /                   \
                /                     \
             10-v                    10-v
              /                         \
             /                           \
            o                             o
         -(10-v)                       -(10-v)
          Node4                         Node4
         FIGURE 3                      FIGURE 4
```

2.(c) The basic solution corresponding to the spanning tree T_2 in FIGURE 4
has been calculated as follows:
$x_{13} = u$, due to the flow balance requirement in node 1,
$x_{23} = v - u$, due to the flow balance requirement in node 3,
$x_{24} = 10 - v$, due to the flow balance requirement in node 2.
Then the flow balance requirement in node 4 is also fulfilled,
The basic solution corresponding to T_2 is thus $\hat{\mathbf{x}} = (0,\, u,\, v-u,\, 10-v,\, 0)^\mathsf{T}$,
which is illustrated in FIGURE 4.
If $0 < v < u < 10$ then $\hat{\mathbf{x}}$ is *not feasible*, since $\hat{x}_{23} = v - u < 0$, and thus *not optimal*,.
If $0 < u < v < 10$ then $\hat{\mathbf{x}}$ is a BFS, since $\hat{\mathbf{x}} \geq \mathbf{0}$. It remains to show that it is optimal.
The simplex multipliers y_i for the nodes are calculated by $y_4 = 0$ and $y_i - y_j = c_{ij}$ for all arcs (i,j) in the spanning tree. Using that $c_{ij} = 1$ for all arcs, the y_i are calculated in the order
$y_4 = 0,\; y_2 = y_4 + c_{24} = 1,\; y_3 = y_2 - c_{23} = 0,\; y_1 = y_3 + c_{23} = 1.$
Then the reduced cost for the two non-basic variables are calculated by $r_{ij} = c_{ij} - y_i + y_j$, which give that $r_{12} = c_{12} - y_1 + y_2 = 1$ and $r_{34} = c_{34} - y_3 + y_4 = 1$.
Since both $r_{12} > 0$ and $r_{34} > 0$, the current BFS $\hat{\mathbf{x}}$ is a unique optimal solution.

3.(a) When $C = 0$, the considered problem can be written

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

where $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 \\ 3 & 1 & -4 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$.

We use elementary row operations (Gauss-Jordan) to put the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ on reduced row echelon form: $\left[\begin{array}{ccc|c} 1 & 2 & -3 & 10 \\ 3 & 1 & -4 & 10 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 2 \\ 0 & 1 & -1 & 4 \end{array}\right]$

The general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is then obtained by letting $x_3 = v$ (an arbitrary number) whereafter $x_1 = 2 + v$ and $x_2 = 4 + v$.
Thus, the complete set of solutions to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is given by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} v = \bar{\mathbf{x}} + \mathbf{z}\, v,$$

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

Changing variables from $\mathbf{x}$ to v leads to a quadratic objective function which is minimized by any solution 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 semidefinite (≥ 0 in this one-variable case) and at least one such solution v exists.

We get that $\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z} = \mathbf{z}^{\mathsf{T}}\mathbf{z} = 3 > 0$ and $-\mathbf{z}^{\mathsf{T}}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}) = -\mathbf{z}^{\mathsf{T}}\bar{\mathbf{x}} = -6$,

so the unique solution to the above system is $\hat{v} = -6/3 = -2$,

and the unique optimal solution to the original problem P is

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}, \text{ with } f(\hat{\mathbf{x}}) = 4.$$

The Lagrange optimality conditions for the considered problem P are given by the system $\begin{array}{rcr} \mathbf{H}\mathbf{x} - \mathbf{A}^{\mathsf{T}}\mathbf{u} & = & -\mathbf{c} \\ \mathbf{A}\mathbf{x} & = & \mathbf{b} \end{array}$

The above optimal vector $\hat{\mathbf{x}}$ of course satisfies $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$.

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

which when $\mathbf{x} = \hat{\mathbf{x}}$ becomes $\begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ -3 & -4 \end{bmatrix} \begin{pmatrix} u_1 \\ u_1 \end{pmatrix}$.

Solving the first two of these three equations gives the unique solution $\hat{u}_1 = 1.2$, $\hat{u}_2 = -0.4$, which also satisfies the third equation.

3.(b) and 3.(c)

When $C \neq 0$, the matrix $\mathbf{H}$ becomes $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -C \\ 0 & -C & 1 \end{bmatrix}$,

while $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\bar{\mathbf{x}} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}$ and $\mathbf{z} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ are the same as above.

Then $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z} = 3-2C$ and $-\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}) = -\mathbf{z}^\mathsf{T}\mathbf{H}\bar{\mathbf{x}} = -(6-4C)$.

From this, the following conclusions can be drawn:

If $C > 1.5$ then $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z}$ is not positive semidefinite, and then the considered problem has no optimal solution.

If $C < 1.5$ then $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z}$ is positive definite, and the system $(\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z})\, v = -\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, i.e. the equation $(3-C)\, v = -(6-4C)$, has the *unique* solution $\hat{v} = -2$.
Then the original problem P has the *unique* optimal solution

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}, \text{ with } f(\hat{\mathbf{x}}) = 4 + 4C.$$

If $C = 1.5$ then $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z}$ is positive semidefinite but not positive definite, and the system $(\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z})\, v = -\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$, i.e. the equation $(3-C)\, v = -(6-4C)$, becomes $0\, v = 0$, which is satisfied for all $v \in \mathbb{R}$.

Then the original problem P has an infinite number of optimal solutions, namely all the vectors

$$\mathbf{x}(v) = \bar{\mathbf{x}} + \mathbf{z}\, v = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} + \begin{pmatrix} v \\ v \\ v \end{pmatrix} = \begin{pmatrix} 2+v \\ 4+v \\ 0+v \end{pmatrix}, \text{ where } v \in \mathbb{R},$$

with $f(\mathbf{x}(v)) = \frac{1}{2}\left((2+v)^2 + (4+v)^2 + (0+v)^2 - 3(4+v)(0+v)\right) = 10.$
Thus, when $C = 1.5$, the set of optimal solutions to P is a line in $\mathbb{R}^3$.

Answer on 3.(b): $C < 1.5$, $\hat{\mathbf{x}} = (0\,,\,2\,,-2)^\mathsf{T}$, $f(\hat{\mathbf{x}}) = 4 + 4C$.

Answer on 3.(c): $C = 1.5$, $\mathbf{x}(v) = (2+v\,,\,4+v\,,\,v)^\mathsf{T}$ for $v \in \mathbb{R}$, $f(\mathbf{x}(v)) = 10$.

4.(a) Some calculations give that

$h_1(\mathbf{x})+h_2(\mathbf{x})+h_3(\mathbf{x})+h_4(\mathbf{x}) = 4x_1^2 + 4x_2^2 - 4x_2 + 4 = \frac{1}{2}\,\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} + c_0\,,$

with $\mathbf{H} = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 0 \\ -4 \end{pmatrix}$ and $c_0 = 4$.

Since $\mathbf{H}$ is a diagonal matrix with strictly positive diagonal elements, $\mathbf{H}$ is positive definite. This implies that a unique global minimum point to the above quadratic function is obtain by solving $\mathbf{H}\mathbf{x} = -\mathbf{c}$, which has the solution $\tilde{\mathbf{x}} = (0\,,0.5)^\mathsf{T}$, with $\frac{1}{2}\,\tilde{\mathbf{x}}^\mathsf{T}\mathbf{H}\tilde{\mathbf{x}} + \mathbf{c}^\mathsf{T}\tilde{\mathbf{x}} + c_0 = 3$.
Thus, we can conclude that $h_1(\mathbf{x})+h_2(\mathbf{x})+h_3(\mathbf{x})+h_4(\mathbf{x}) \geq 3 > 0$ for all $\mathbf{x} \in I\!R^2$.
Now, if a point $\bar{\mathbf{x}} \in I\!R^2$ satisfies that $h_i(\bar{\mathbf{x}}) = 0,\ i = 1,2,3,4$, then
$h_1(\bar{\mathbf{x}})+h_2(\bar{\mathbf{x}})+h_3(\bar{\mathbf{x}})+h_4(\bar{\mathbf{x}}) = 0+0+0+0 = 0$, which is impossible according to above.
Thus, we can conclude that there is no solution to the system $h_i(\mathbf{x}) = 0,\ i = 1,2,3,4$.

4.(b) Let $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), h_2(\mathbf{x}), h_3(\mathbf{x}), h_4(\mathbf{x}))^\mathsf{T}$.

Then $f(\mathbf{x}) = \frac{1}{2}\,(\,h_1(\mathbf{x})^2 + h_2(\mathbf{x})^2 + h_3(\mathbf{x})^2 + h_4(\mathbf{x})^2) = \frac{1}{2}\,\mathbf{h}(\mathbf{x})^\mathsf{T}\mathbf{h}(\mathbf{x})$.

Since f has continuous derivatives, a necessary condition for a point $\hat{\mathbf{x}} \in I\!R^2$ to be a local minimum point to $f(\mathbf{x})$ (without any constraint) is that $\nabla f(\hat{\mathbf{x}}) = \mathbf{0}^\mathsf{T}$.

By the chain rule, the gradient of f is given by

$$\nabla f(\hat{\mathbf{x}}) = \mathbf{h}(\mathbf{x})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}), \text{ where } \nabla\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 2(x_1+2) & 2(x_2+1) \\ 2(x_1-2) & 2(x_2+1) \\ 2(x_1+1) & 2(x_2-2) \\ 2(x_1-1) & 2(x_2-2) \end{bmatrix}.$$

$$\text{In particular, } \nabla f(\mathbf{0}) = \mathbf{h}(\mathbf{0})^\mathsf{T}\nabla\mathbf{h}(\mathbf{0}) = (1\,,1\,,1\,,1)\begin{bmatrix} 4 & 2 \\ -4 & 2 \\ 2 & -4 \\ -2 & -4 \end{bmatrix} = (0\,,-4) \neq \mathbf{0}^\mathsf{T}.$$

Thus, $\mathbf{x} = \mathbf{0}$ is not even a local optimal solution to the considered least squares problem.

4.(c) Since f has continuous second derivatives, f is a convex function on $I\!R^2$ if and only if its Hessian matrix $\mathbf{F}(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in I\!R^2$.

By the chain rule, the Hessian is given by $\mathbf{F}(\mathbf{x}) = \nabla\mathbf{h}(\mathbf{x})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}) + \sum_i h_i(\mathbf{x})\mathbf{H}_i(\mathbf{x})$.

But $\mathbf{H}_i(\mathbf{x}) = 2\,\mathbf{I}$ for all i. Therefore, $\sum_i h_i(\mathbf{x})\mathbf{H}_i(\mathbf{x}) = (2\sum_i h_i(\mathbf{x}))\,\mathbf{I}$.

Let $\mathbf{x}$ be an arbitrary point in $I\!R^2$. We shall check if $\mathbf{F}(\mathbf{x})$ is positive semidefinite.

For any vector $\mathbf{w} \in I\!R^2$ we get that

$\mathbf{w}^\mathsf{T}\mathbf{F}(\mathbf{x})\mathbf{w} = \mathbf{w}^\mathsf{T}(\nabla\mathbf{h}(\mathbf{x})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}) + (2\sum_i h_i(\mathbf{x}))\,\mathbf{I}\,)\mathbf{w} =$

$\mathbf{w}^\mathsf{T}\nabla\mathbf{h}(\mathbf{x})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x})\mathbf{w} + \mathbf{w}^\mathsf{T}\,(2\sum_i h_i(\mathbf{x}))\,\mathbf{I}\,\mathbf{w} = (\nabla\mathbf{h}(\mathbf{x})\mathbf{w})^\mathsf{T}(\nabla\mathbf{h}(\mathbf{x})\mathbf{w}) + (2\sum_i h_i(\mathbf{x}))\,\mathbf{w}^\mathsf{T}\mathbf{w} =$

$\|\,\nabla\mathbf{h}(\mathbf{x})\mathbf{w}\,\|^2 + (2\sum_i h_i(\mathbf{x}))\,\|\,\mathbf{w}\,\|^2 \geq 0$, since $\sum_i h_i(\mathbf{x}) > 0$ according to 4.(a).

This shows that $\mathbf{F}(\mathbf{x})$ is positive semidefinite in the arbitrary point $\mathbf{x} \in I\!R^2$,
which means that $\mathbf{F}(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in I\!R^2$,
which implies that f is a convex function on $I\!R^2$.

5. The Lagrange function for the problem can be written

$$L(\mathbf{x},\mathbf{y}) = \tfrac{1}{2}\,(x_1-q_1)^2 + \tfrac{1}{2}\,(x_2-q_2)^2 + \tfrac{1}{2}\,(x_3-q_3)^2 + \sum_{i=1}^{8} y_i g_i(\mathbf{x}),$$

where

$$\begin{aligned}
g_1(\mathbf{x}) &= x_1 + x_2 + x_3 - 1,\\
g_2(\mathbf{x}) &= x_1 + x_2 - x_3 - 1,\\
g_3(\mathbf{x}) &= x_1 - x_2 + x_3 - 1,\\
g_4(\mathbf{x}) &= x_1 - x_2 - x_3 - 1,\\
g_5(\mathbf{x}) &= -x_1 + x_2 + x_3 - 1,\\
g_6(\mathbf{x}) &= -x_1 + x_2 - x_3 - 1,\\
g_7(\mathbf{x}) &= -x_1 - x_2 + x_3 - 1,\\
g_8(\mathbf{x}) &= -x_1 - x_2 - x_3 - 1.
\end{aligned}$$

For fixed $\mathbf{y}$, this Lagrange function is a strictly convex quadratic function in $\mathbf{x}$. Therefore, the unique $\mathbf{x}$ which minimizes $L(\mathbf{x},\mathbf{y})$, for fixed $\mathbf{y}$, is obtained by setting the partial derivatives of L with respect the primal variables x_j equal to zero.

The global optimality conditions (GOC) then becomes:

$$\begin{aligned}
\hat{x}_1 - q_1 + \hat{y}_1 + \hat{y}_2 + \hat{y}_3 + \hat{y}_4 - \hat{y}_5 - \hat{y}_6 - \hat{y}_7 - \hat{y}_8 &= 0,\\
\hat{x}_2 - q_2 + \hat{y}_1 + \hat{y}_2 - \hat{y}_3 - \hat{y}_4 + \hat{y}_5 + \hat{y}_6 - \hat{y}_7 - \hat{y}_8 &= 0, \qquad \text{(GOC-1)}\\
\hat{x}_3 - q_3 + \hat{y}_1 - \hat{y}_2 + \hat{y}_3 - \hat{y}_4 + \hat{y}_5 - \hat{y}_6 + \hat{y}_7 - \hat{y}_8 &= 0,
\end{aligned}$$

$$\begin{aligned}
&g_i(\hat{\mathbf{x}}) \le 0, && \text{for } i = 1,\dots,8, && \text{(GOC-2)}\\
&\hat{y}_i \ge 0, && \text{for } i = 1,\dots,8, && \text{(GOC-3)}\\
&\hat{y}_i g_i(\hat{\mathbf{x}}) = 0, && \text{for } i = 1,\dots,8. && \text{(GOC-4)}
\end{aligned}$$

5.(a) Assume that $\mathbf{q} = (-0.5,\ 0.4, -0.4)^{\mathsf{T}}$.

If $\hat{\mathbf{x}} = (-0.4,\ 0.3, -0.3)^{\mathsf{T}}$ then $g_i(\hat{\mathbf{x}}) = 0$ for $i = 6$, while $g_i(\hat{\mathbf{x}}) < 0$ for all $i \neq 6$.

Thus, (GOC-2) is satisfied. Further, (GOC-4) implies that $\hat{y}_i = 0$ for all $i \neq 6$.

Then (GOC-1) becomes:

$$\begin{aligned}
-0.4 + 0.5 - \hat{y}_6 &= 0,\\
0.3 - 0.4 + \hat{y}_6 &= 0,\\
-0.3 + 0.4 - \hat{y}_6 &= 0,
\end{aligned}$$

which is satisfied by $\hat{y}_6 = 0.1$.

Since $\hat{y}_6 > 0$, (GOC-3) is also satisfied.

Thus, $\hat{\mathbf{x}} = (-0.4,\ 0.3, -0.3)^{\mathsf{T}}$, together with $\hat{\mathbf{y}} = (0,0,0,0,0,0.1,0,0)^{\mathsf{T}}$, satisfies all the global optimality conditions. By a well-known result, this implies that $\hat{\mathbf{x}} = (-0.4,\ 0.3, -0.3)^{\mathsf{T}}$ is an optimal solution to P when $\mathbf{q} = (-0.5,\ 0.4, -0.4)^{\mathsf{T}}$.

5.(b) Assume that $\mathbf{q} = (-0.8,\ 0.6, -0.1)^{\mathsf{T}}$.

If $\hat{\mathbf{x}} = (-0.6,\ 0.4,\ 0)^{\mathsf{T}}$ then $g_i(\hat{\mathbf{x}}) = 0$ for $i{=}5$ and $i{=}6$, while $g_i(\hat{\mathbf{x}}) < 0$ for all other i.

Thus, (GOC-2) is satisfied. Further, (GOC-4) implies that $\hat{y}_i = 0$ for $i = 1, 2, 3, 4, 7, 8$.

Then (GOC-1) becomes:

$$\begin{aligned} -0.6 + 0.8 - \hat{y}_5 - \hat{y}_6 &= 0, \\ 0.4 - 0.6 + \hat{y}_5 + \hat{y}_6 &= 0, \\ 0.0 + 0.1 + \hat{y}_5 - \hat{y}_6 &= 0, \end{aligned}$$

which is satisfied by $\hat{y}_6 = 0.05$ and $\hat{y}_6 = 0.15$.

Since $\hat{y}_5 > 0$ and $\hat{y}_6 > 0$, (GOC-3) is also satisfied.

Thus, $\hat{\mathbf{x}} = (-0.6,\ 0.4,\ 0)^{\mathsf{T}}$, together with $\hat{\mathbf{y}} = (0, 0, 0, 0, 0.05, 0.15, 0, 0)^{\mathsf{T}}$, satisfies all the global optimality conditions. By a well-known result, this implies that $\hat{\mathbf{x}} = (-0.6,\ 0.4,\ 0)^{\mathsf{T}}$ is an optimal solution to P when $\mathbf{q} = (-0.8,\ 0.6, -0.1)^{\mathsf{T}}$.