Solutions to exam in SF1811 Optimization, Jan 14, 2015

```
                  1      3
              300 O------O -400
                   \    /
                    \  /
The network:         \/     where all links go from left to right.
                     /\
                    /  \
                   /    \
              600 O------O -500
                  2      4
```

1.(a) Let $\mathbf{x} = (\, x_{13},\ x_{14},\ x_{23},\ x_{24}\,)^\mathsf{T}$, where the variable x_{ij} stands for the flow in the link from node i to node j, and let $\mathbf{c} = (\, c_{13},\ c_{14},\ c_{23},\ c_{24}\,)^\mathsf{T}$.
Then the total cost for the flow is given by $\mathbf{c}^\mathsf{T}\mathbf{x}$.
The flow balance conditions in the nodes can be written $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 300 \\ 600 \\ -400 \\ -500 \end{pmatrix}.$$

Finally, the given directions of the links imply the constraints $\mathbf{x} \geq \mathbf{0}$.
(It is recommended to remove the last row in $\mathbf{A}$ and the corresponding last component in $\mathbf{b}$ to get a system without any redundant equation.)

1.(b) The four different spanning trees are shown in the following figure, together wih the unique link flows which satisfy $\mathbf{A}\mathbf{x} = \mathbf{b}$.

```
   -200          400           300
 O------O      O------O      O------O      O      O
  \    /        \                  /        \    /
500\  /          \                /      300\  /
    \/       -100\               /           \/
    /\            \          100/            /\
600/  \            \           /         400/  \
  /    \            \         /            /    \
 O      O      O------O      O------O      O------O
                 600           500           200
 Spanning      Spanning      Spanning      Spanning
 tree 1        tree 2        tree 3        tree 4
```

The link flows in the first spanning tree are calculated as follows:
The only way to satisfy the supply constraint in node 2 is to let $x_{23} = 600$.
Then the only way to satisfy the demand constraint in node 3 is to let $x_{13} = -200$.
Then the only way to satisfy the supply constraint in node 1 is to let $x_{14} = 500$.
Then the demand constraint in node 4 is also satisfied.

The link flows for the other spanning trees are calculated in a similar way.

We thus have the following four basic solutions:
$\mathbf{x} = (-200,\ 500,\ 600,\ 0)^{\mathsf{T}}$, corresponding to spanning tree number 1,
$\mathbf{x} = (\ 400,\ -100,\ 0,\ 600)^{\mathsf{T}}$, corresponding to spanning tree number 2,
$\mathbf{x} = (\ 300,\ 0,\ 100,\ 500)^{\mathsf{T}}$, corresponding to spanning tree number 3, and
$\mathbf{x} = (\ 0,\ 300,\ 400,\ 200)^{\mathsf{T}}$, corresponding to spanning tree number 4.

All these four solutions satisfy $\mathbf{Ax} = \mathbf{b}$, but only the last two satisfy $\mathbf{x} \geq \mathbf{0}$.

Thus, the basic solutions corresponding to spanning trees 3 and 4 are *feasible* basic solutions, while the basic solutions corresponding to spanning trees 1 and 2 are *infeasible* basic solutions.

1.(c) For a given feasible basic solution, the simplex multipliers y_i for the different nodes are calculated from $y_4 = 0$ and $y_i - y_j = c_{ij}$ for all links (i, j) in the corresponding spanning tree.

For the feasible basic solution corresponding to spanning tree 3, we get
$y_4 = 0$,
$y_2 = y_4 + c_{24} = c_{24}$,
$y_3 = y_2 - c_{23} = c_{24} - c_{23}$,
$y_1 = y_3 + c_{13} = c_{24} - c_{23} + c_{13}$.
The reduced cost for the only non-basic variable is then given by
$r_{14} = c_{14} - y_1 + y_4 = c_{14} - c_{24} + c_{23} - c_{13}$.
Thus, if $c_{13} - c_{14} - c_{23} + c_{24} < 0$ then $r_{14} > 0$, and then $\mathbf{x} = (300,\ 0,\ 100,\ 500)^{\mathsf{T}}$
is the unique optimal solution to the considered problem.

For the feasible basic solution corresponding to spanning tree 4, we get
$y_4 = 0$,
$y_2 = y_4 + c_{24} = c_{24}$,
$y_1 = y_4 + c_{14} = c_{14}$,
$y_3 = y_2 - c_{23} = c_{24} - c_{23}$,
The reduced cost for the only non-basic variable is then given by
$r_{13} = c_{13} - y_1 + y_3 = c_{13} - c_{14} + c_{24} - c_{23}$.
Thus, if $c_{13} - c_{14} - c_{23} + c_{24} > 0$ then $r_{13} > 0$, and then $\mathbf{x} = (\ 0,\ 300,\ 400,\ 200)^{\mathsf{T}}$
is the unique optimal solution to the considered problem.

If $c_{13} - c_{14} - c_{23} + c_{24} = 0$ then both the above feasible basic solutions are optimal solutions, and then every convex combination of these two solutions, i.e.
$\mathbf{x} = t\,(300,\ 0,\ 100,\ 500)^{\mathsf{T}} + (1-t)\,(\ 0,\ 300,\ 400,\ 200)^{\mathsf{T}}$, where $t \in [0, 1]$,
is also an optimal solution since the constraints and the objective function are linear.
As an example, the following three solutions (corresponding to $t = 1/3, 1/2$ and $2/3$) are optimal solutions for the case that $c_{13} - c_{14} - c_{23} + c_{24} = 0$:
$\mathbf{x} = (100,\ 200,\ 300,\ 300)^{\mathsf{T}}$, $\mathbf{x} = (150,\ 150,\ 250,\ 350)^{\mathsf{T}}$ and $\mathbf{x} = (200,\ 100,\ 200,\ 400)^{\mathsf{T}}$.

2.(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 $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & 0 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$ and $\mathbf{c}^\mathsf{T} = (4,\ 4,\ 2,\ 4,\ 4)$.

The starting solution should have the basic variables x_1 and x_5, which means that $\beta = (1, 5)$ and $\nu = (2, 3, 4)$.

The corresponding basic matrix is $\mathbf{A}_\beta = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}$, while $\mathbf{A}_\nu = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$.

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.

$$\begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

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

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

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

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

Since $r_{\nu_2} = r_3 = -2$ 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.

$$\begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix} \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{a}}_3 = \begin{pmatrix} \bar{a}_{13} \\ \bar{a}_{23} \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

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

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i3}} \mid \bar{a}_{i3} > 0 \right\} = \min \left\{ \frac{1}{0.5}, \frac{2}{0.5} \right\} = \frac{1}{0.5} = \frac{\bar{b}_1}{\bar{a}_{13}}.$$

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 $\beta = (3, 5)$ and $\nu = (2, 1, 4)$.

The corresponding basic matrix is $\mathbf{A}_\beta = \begin{bmatrix} 2 & 4 \\ 2 & 0 \end{bmatrix}$, while $\mathbf{A}_\nu = \begin{bmatrix} 1 & 0 & 3 \\ 3 & 4 & 1 \end{bmatrix}$.

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.

$$\begin{bmatrix} 2 & 4 \\ 2 & 0 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

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

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

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

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

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

Thus, $\mathbf{x} = (0, 0, 2, 0, 1)^\mathsf{T}$ is an optimal solution, with optimal value $\mathbf{c}^\mathsf{T}\mathbf{x} = 8$.

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

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

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

$$\begin{array}{lrcrcl} \text{maximize} & 8y_1 & + & 4y_2 & & \\ \text{subject to} & & & 4y_2 & \leq & 4, \\ & y_1 & + & 3y_2 & \leq & 4, \\ & 2y_1 & + & 2y_2 & \leq & 2, \\ & 3y_1 & + & y_2 & \leq & 4, \\ & 4y_1 & & & \leq & 4. \end{array}$$

This dual problem can be illustrated by drawing the constraints and some level lines for the objective function in a coordinate system with y_1 and y_2 on the axes. (The figure is omitted here.)

It is well known that an optimal solution to this problem is given by the vector $\mathbf{y}$ of "simplex multipliers" for the optimal basic solution in (a) above, i.e. $\mathbf{y} = (1,\ 0)^\mathsf{T}$. Alternatively, this can be seen from the figure (which is omitted here).

Check: It is easy to verify that $\mathbf{y} = (1,\ 0)^\mathsf{T}$ satisfies the dual constraints, with dual objective value $8y_1 + 4y_2 = 8 =$ the optimal value of the primal problem. Thus, $\mathbf{y} = (1,\ 0)^\mathsf{T}$ is an optimal solution to the dual problem.

2.(c)

If the second constraint in the primal problem is removed, the corresponding dual problem becomes

$$\begin{array}{rrcl} \text{maximize} & 8y & & \\ \text{subject to} & 0y & \leq & 4, \\ & y & \leq & 4, \\ & 2y & \leq & 2, \\ & 3y & \leq & 4, \\ & 4y & \leq & 4. \end{array}$$

The optimal solution to this problem is clearly $y = 1$, with the optimal value $8y = 8$.
But then the optimal value of the reduced primal problem must also be $= 8$.
Since the optimal solution $\mathbf{x} = (0, 0, 2, 0, 1)^\mathsf{T}$ from (a) above is feasible also to the reduced primal problem, and still has the objective value $\mathbf{c}^\mathsf{T}\mathbf{x} = 8$, it follows that $\mathbf{x} = (0, 0, 2, 0, 1)^\mathsf{T}$ is an optimal solution also to the reduced primal problem!
(But not a basic solution. Two optimal basic solutions are now $\mathbf{x} = (0, 0, 4, 0, 0)^\mathsf{T}$ and $\mathbf{x} = (0, 0, 0, 0, 2)^\mathsf{T}$, with objective values $= 8$.)

2.(d)

If the first constraint in the primal problem is removed, the corresponding dual problem becomes

$$\begin{array}{rrcl} \text{maximize} & 4y & & \\ \text{subject to} & 4y & \leq & 4, \\ & 3y & \leq & 4, \\ & 2y & \leq & 2, \\ & y & \leq & 4, \\ & 0y & \leq & 4. \end{array}$$

The optimal solution to this problem is clearly $y = 1$, with the optimal value $4y = 4$,
and then the optimal value of the reduced primal problem must also be $= 4$.
But the optimal solution $\mathbf{x} = (0, 0, 2, 0, 1)^\mathsf{T}$ from (a) above still has the objective value $\mathbf{c}^\mathsf{T}\mathbf{x} = 8 > 4$, so it can *not* be an optimal solution to the reduced primal problem!
(Two optimal basic solutions are now $\mathbf{x} = (0, 0, 2, 0, 0)^\mathsf{T}$ and $\mathbf{x} = (1, 0, 0, 0, 0)^\mathsf{T}$,
with objective values $= 4$.)

3.(a)

The objective function is $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$, with $\mathbf{H} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 1 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

LDL^T-factorization of $\mathbf{H}$ gives

$$\mathbf{H} = \mathbf{LDL}^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Since there is a negative diagonal element in $\mathbf{D}$, the matrix $\mathbf{H}$ is *not* positive semidefinite, which in turn implies that there is no optimal solution to the problemen of minimizing $f(\mathbf{x})$ without constraints. (With e.g. $\mathbf{d} = (1,1,1)^\mathsf{T}$, $f(t\,\mathbf{d}) = -t^2 \to -\infty$ when $t \to \infty$.)

3.(b)

We now 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 $\mathbf{H}$ and $\mathbf{c}$ as above, $\mathbf{A} = [\,1 \;\; -1 \;\;\; 1\,]$ and $\mathbf{b} = 0$.

The general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, i.e. to $x_1 - x_2 + x_3 = 0$, is given by

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

which means that $\bar{\mathbf{x}} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is a feasible solution, and $\mathbf{Z} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ is a matrix whos columns form a basis for the null space 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 $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z} = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$, which is positive semidefinite (but not positive definite).

The system $(\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z})\mathbf{v} = -\mathbf{Z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ becomes $\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,

with the solutions $\hat{\mathbf{v}}(t) = \begin{pmatrix} 2t \\ t \end{pmatrix}$, for arbitrary values on the real number t, which implies that $\hat{\mathbf{x}}(t) = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}}(t) = \begin{pmatrix} t \\ 2t \\ t \end{pmatrix}$, for $t \in I\!R$, are the (infinite number of) optimal solutions.

Note that $f(\hat{\mathbf{x}}(t)) = 0$ for all $t \in I\!R$.

3.(c)

Again, we have a problem on 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 $\mathbf{H}$ and $\mathbf{c}$ as above, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, and $\mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

The general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is now $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \cdot v$, for $v \in I\!R$, which implies that

$\bar{\mathbf{x}} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is a feasible solution, and $\mathbf{z} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ form a basis for the null space of $\mathbf{A}$.

After the variable change $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{z}v$, we should solve 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 at least ≥ 0.
We have that $\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z} = 6 > 0$, so the system $(\mathbf{z}^\mathsf{T}\mathbf{H}\mathbf{z})v = -\mathbf{z}^\mathsf{T}(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ becomes $6v = 0$, with the unique solution $\hat{v} = 0$, so that $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = \mathbf{0}$ is the unique optimal solution.

4.(a) The Lagrange function for the considered problem is given by

$L(\mathbf{x},\mathbf{y}) = \frac{1}{2}\,(\mathbf{x}-\mathbf{q})^{\mathsf{T}}(\mathbf{x}-\mathbf{q}) + \mathbf{y}^{\mathsf{T}}(\mathbf{b}-\mathbf{A}\mathbf{x})$, with $\mathbf{x} \in I\!R^n$ and $y \in I\!R^m$.

The Lagrange relaxed problem $\mathrm{PR}_{\mathbf{y}}$ is defined, for a given $\mathbf{y} \geq \mathbf{0}$, as the problem of minimizing $L(\mathbf{x},\mathbf{y})$ with respect to $\mathbf{x} \in I\!R^n$.

Since $L(\mathbf{x},\mathbf{y}) = \frac{1}{2}\,\mathbf{x}^{\mathsf{T}}\mathbf{I}\,\mathbf{x} - (\mathbf{A}^{\mathsf{T}}\mathbf{y}+\mathbf{q})^{\mathsf{T}}\mathbf{x} + \mathbf{b}^{\mathsf{T}}\mathbf{y} + \frac{1}{2}\,\mathbf{q}^{\mathsf{T}}\mathbf{q}\,,$
and the unit matrix $\mathbf{I}$ is positive definite, the unique optimal solution to $\mathrm{PR}_{\mathbf{y}}$ is given by $\tilde{\mathbf{x}}(\mathbf{y}) = \mathbf{A}^{\mathsf{T}}\mathbf{y} + \mathbf{q}$.

Then the dual objective function becomes

$$\begin{aligned}\varphi(\mathbf{y}) = L(\tilde{\mathbf{x}}(\mathbf{y}),\mathbf{y}) &= -\tfrac{1}{2}\,(\mathbf{A}^{\mathsf{T}}\mathbf{y}+\mathbf{q})^{\mathsf{T}}(\mathbf{A}^{\mathsf{T}}\mathbf{y}+\mathbf{q}) + \mathbf{b}^{\mathsf{T}}\mathbf{y} + \tfrac{1}{2}\,\mathbf{q}^{\mathsf{T}}\mathbf{q} = \\ &= -\tfrac{1}{2}\,\mathbf{y}^{\mathsf{T}}\mathbf{A}\mathbf{A}^{\mathsf{T}}\mathbf{y} + (\mathbf{b}-\mathbf{A}\mathbf{q})^{\mathsf{T}}\mathbf{y}.\end{aligned}$$

4.(b) From now on, $\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$ and $\mathbf{q} = (1,2,2,1)^{\mathsf{T}}$.

Then $\mathbf{A}\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ and $\mathbf{b}-\mathbf{A}\mathbf{q} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$,

so that the dual objective function becomes $\varphi(\mathbf{y}) = -2y_1^2 - 2y_2^2 + 4y_1 - y_2\,.$

Alternative calculation of the dual function, without using the results from (a):

The considered problem is: minimize $f(\mathbf{x})$ subject to $g_1(\mathbf{x}) \leq 0$ and $g_2(\mathbf{x}) \leq 0$,

where $f(\mathbf{x}) = \frac{1}{2}\,(x_1-1)^2 + \frac{1}{2}\,(x_2-2)^2 + \frac{1}{2}\,(x_3-2)^2 + \frac{1}{2}\,(x_4-1)^2$,
$g_1(\mathbf{x}) = 6 - x_1 - x_2 + x_3 - x_4$ and $g_2(\mathbf{x}) = 3 - x_1 - x_2 - x_3 + x_4$.

The Lagrange function then becomes:

$$\begin{aligned}L(\mathbf{x},\mathbf{y}) = {} & \tfrac{1}{2}\,(x_1-1)^2 + \tfrac{1}{2}\,(x_2-2)^2 + \tfrac{1}{2}\,(x_3-2)^2 + \tfrac{1}{2}\,(x_4-1)^2 + \\ & y_1\,(6 - x_1 - x_2 + x_3 - x_4) + y_2\,(3 - x_1 - x_2 - x_3 + x_4) = \\ = {} & \tfrac{1}{2}\,(x_1-1)^2 - (y_1{+}y_2)\,x_1 + \tfrac{1}{2}\,(x_2-2)^2 - (y_1{+}y_2)\,x_2 + \\ & \tfrac{1}{2}\,(x_3-2)^2 - (y_2{-}y_1)\,x_3 + \tfrac{1}{2}\,(x_4-1)^2 - (y_1{-}y_2)\,x_4 + 6y_1 + 3y_2.\end{aligned}$$

Minimizing this with respect to $\mathbf{x}$ gives:

$\tilde{x}_1(\mathbf{y}) = 1{+}y_1{+}y_2,\ \ \tilde{x}_2(\mathbf{y}) = 2{+}y_1{+}y_2,\ \ \tilde{x}_3(\mathbf{y}) = 2{+}y_2{-}y_1,\ \ \tilde{x}_4(\mathbf{y}) = 1{+}y_1{-}y_2,$

so that $\tilde{\mathbf{x}}(\mathbf{y}) = (1{+}y_1{+}y_2,\ 2{+}y_1{+}y_2,\ 2{+}y_2{-}y_1,\ 1{+}y_1{-}y_2)^{\mathsf{T}}$,

and then the dual function becomes $\varphi(\mathbf{y}) = L(\tilde{\mathbf{x}}(\mathbf{y}),\mathbf{y}) =$

$$\begin{aligned}& \tfrac{1}{2}\,(\tilde{x}(\mathbf{y})_1 - 1)^2 - (y_1{+}y_2)\,\tilde{x}(\mathbf{y})_1 + \tfrac{1}{2}\,(\tilde{x}(\mathbf{y})_2 - 2)^2 - (y_1{+}y_2)\,\tilde{x}(\mathbf{y})_2 + \\ & \tfrac{1}{2}\,(\tilde{x}(\mathbf{y})_3 - 2)^2 - (y_2{-}y_1)\,\tilde{x}(\mathbf{y})_3 + \tfrac{1}{2}\,(\tilde{x}(\mathbf{y})_4 - 1)^2 - (y_1{-}y_2)\,\tilde{x}(\mathbf{y})_4 + 6y_1 + 3y_2 \ = \\ & \tfrac{1}{2}\,(y_1{+}y_2)^2 - (y_1{+}y_2) - (y_1{+}y_2)^2 + \tfrac{1}{2}\,(y_1{+}y_2)^2 - 2(y_1{+}y_2) - (y_1{+}y_2)^2 + \\ & \tfrac{1}{2}\,(y_2{-}y_2)^2 - 2(y_2{-}y_1) - (y_2{-}y_1)^2 + \tfrac{1}{2}\,(y_1{-}y_2)^2 - (y_1{-}y_2) - (y_1{-}y_2)^2 + \\ & 6y_1 + 3y_2 = \ldots\ldots = -2y_1^2 - 2y_2^2 + 4y_1 - y_2, \text{ as above.}\end{aligned}$$

The dual problem then becomes:

D: maximize $\varphi(\mathbf{y}) = -2y_1^2 - 2y_2^2 + 4y_1 - y_2$ subject to $y_1 \geq 0$ and $y_2 \geq 0$,

which decomposes into the two separate problems

D_1: maximize $-2y_1^2 + 4y_1$ subject to $y_1 \geq 0$, and

D_2: maximize $-2y_2^2 - y_2$ subject to $y_2 \geq 0$.

Clearly, the optimal solution to the first problem is $\hat{y}_1 = 1$,
while the optimal solution to the second problem is $\hat{y}_2 = 0$.

Thus, $\hat{\mathbf{y}} = (1, 0)^\mathsf{T}$ is the unique optimal solution to D, with $\varphi(\hat{\mathbf{y}}) = 2$.

4.(c) Let $\hat{\mathbf{x}} = \tilde{\mathbf{x}}(\hat{\mathbf{y}}) = (1+\hat{y}_1+\hat{y}_2,\ 2+\hat{y}_1+\hat{y}_2,\ 2+\hat{y}_2-\hat{y}_1,\ 1+\hat{y}_1-\hat{y}_2)^\mathsf{T} = (2, 3, 1, 2)^\mathsf{T}$.

Then $\mathbf{A}\hat{\mathbf{x}} - \mathbf{b} = (6, 4)^\mathsf{T} - (6, 3)^\mathsf{T} \geq \mathbf{0}$, so $\hat{\mathbf{x}}$ is a feasible solution to the primal problem.

Further, the primal objective value is $f(\hat{\mathbf{x}}) = \frac{1}{2}\,(\hat{\mathbf{x}}-\mathbf{q})^\mathsf{T}(\hat{\mathbf{x}}-\mathbf{q}) = \frac{1}{2}\,\|(1, 1, -1, 1)^\mathsf{T}\|^2 = 2$.

Since $\hat{\mathbf{x}}$ is feasible to P and $f(\hat{\mathbf{x}}) = \varphi(\hat{\mathbf{y}})$, we conclude that $\hat{\mathbf{x}}$ is an optimal solution to P.

4.(d) Since $\nabla f(\mathbf{x})^\mathsf{T} = \begin{pmatrix} x_1 - 1 \\ x_2 - 2 \\ x_3 - 2 \\ x_4 - 1 \end{pmatrix}$, $\nabla g_1(\mathbf{x})^\mathsf{T} = \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $\nabla g_2(\mathbf{x})^\mathsf{T} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$,

the KKT conditions become:

(KKT-1): $\begin{pmatrix} x_1 - 1 \\ x_2 - 2 \\ x_3 - 2 \\ x_4 - 1 \end{pmatrix} + y_1 \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix} + y_2 \cdot \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

(KKT-2): $6 - x_1 - x_2 + x_3 - x_4 \leq 0$ and $3 - x_1 - x_2 - x_3 + x_4 \leq 0$.

(KKT-3): $y_1 \geq 0$ and $y_2 \geq 0$.

(KKT-4): $y_1\,(6 - x_1 - x_2 + x_3 - x_4) = 0$ and $y_2\,(3 - x_1 - x_2 - x_3 + x_4) = 0$.

With $\mathbf{x} = \hat{\mathbf{x}} = (2, 3, 1, 2)^\mathsf{T}$ and $\mathbf{y} = \hat{\mathbf{y}} = (1, 0)^\mathsf{T}$, we get that

(KKT-1): $\begin{pmatrix} 2 - 1 \\ 3 - 2 \\ 1 - 2 \\ 2 - 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix} + 0 \cdot \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. OK!

(KKT-2): $6 - 2 - 3 + 1 - 2 = 0 \leq 0$ and $3 - 2 - 3 - 1 + 2 = -1 \leq 0$. OK!

(KKT-3): $1 \geq 0$ and $0 \geq 0$. OK!

(KKT-4): $1 \cdot 0 = 0$ and $0 \cdot (-1) = 0$. OK!

Thus, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ satisfy the KKT conditions, and since the considered problem is
a convex problem, we can conclude (again) that $\hat{\mathbf{x}}$ is a global optimal solution.

5.(a)

Change notation and let the variable vector be called $\mathbf{x}$, i.e.

$\mathbf{x} = (x_1, x_2, x_3)^\mathsf{T} = (x, y, r)^\mathsf{T}$.

Then $f(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{m} h_i(\mathbf{x})^2 = \frac{1}{2}\mathbf{h}(\mathbf{x})^\mathsf{T}\mathbf{h}(\mathbf{x})$, where

$h_i(\mathbf{x}) = \sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2} - x_3$ and $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \ldots, h_m(\mathbf{x}))^\mathsf{T}$.

The gradient of h_i is given by

$$\nabla h_i(\mathbf{x}) = \left(\frac{x_1 - a_i}{\sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2}}\,,\ \frac{x_2 - b_i}{\sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2}}\,,\ -1 \right),$$

and $\nabla\mathbf{h}(\mathbf{x})$ is the $m \times 3$ matrix with these gradients as rows.

With the given data, we get that $f(\mathbf{x}) = \frac{1}{2}(h_1(\mathbf{x})^2 + h_2(\mathbf{x})^2 + h_3(\mathbf{x})^2 + h_4(\mathbf{x})^2)$, where

$$\begin{aligned} h_1(\mathbf{x}) &= \sqrt{(x_1 - 5)^2 + x_2^2} - x_3, \\ h_2(\mathbf{x}) &= \sqrt{x_1^2 + (x_2 - 6)^2} - x_3, \\ h_3(\mathbf{x}) &= \sqrt{(x_1 + 4)^2 + x_2^2} - x_3, \\ h_4(\mathbf{x}) &= \sqrt{x_1^2 + (x_2 + 5)^2} - x_3. \end{aligned}$$

The starting point should be $\mathbf{x}^{(1)} = (0,\ 0,\ 5)^\mathsf{T}$. and then

$$\mathbf{h}(\mathbf{x}^{(1)}) = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix},\quad f(\mathbf{x}^{(1)}) = 1 \text{ and } \nabla\mathbf{h}(\mathbf{x}^{(1)}) = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

In Gauss-Newtons method, we should solve $\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\nabla\mathbf{h}(\mathbf{x}^{(1)})\mathbf{d} = -\nabla\mathbf{h}(\mathbf{x}^{(1)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(1)})$,

which becomes $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, with the solution $\mathbf{d}^{(1)} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0 \end{pmatrix}$.

We try $t_1 = 1$, so that $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + t_1\mathbf{d}^{(1)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = (0.5,\ 0.5,\ 5)^\mathsf{T}$. Then

$\mathbf{h}(\mathbf{x}^{(2)}) = (\sqrt{4.5^2 + 0.5^2} - 5,\ \sqrt{5.5^2 + 0.5^2} - 5,\ \sqrt{4.5^2 + 0.5^2} - 5,\ \sqrt{5.5^2 + 0.5^2} - 5)^\mathsf{T} =$

$\frac{1}{2}(\sqrt{82} - 10,\ \sqrt{122} - 10,\ \sqrt{82} - 10,\ \sqrt{122} - 10)^\mathsf{T} \approx \frac{1}{2}(-1,\ 1, -1,\ 1)^\mathsf{T}$, so that

$f(\mathbf{x}^{(2)}) = \frac{1}{2}\mathbf{h}(\mathbf{x}^{(2)})^\mathsf{T}\mathbf{h}(\mathbf{x}^{(2)}) \approx \frac{1}{8}(1 + 1 + 1 + 1) = \frac{1}{2} < 1 = f(\mathbf{x}^{(1)})$,

Thus, the choice $t_1 = 1$ is accepted, and $\mathbf{x}^{(2)} = (0.5,\ 0.5,\ 5)^\mathsf{T}$ is the next iteration point.

5.(b) Let (x_1, x_2) = the location of the (common) center of C_1 and C_2,
z_1 = the square of the radius of the small circle C_1, and
z_2 = the square of the radius of the large circle C_2.

Then the problem can be formulated as follows in the variables are x_1, x_2, z_1 and z_2:

$$\begin{array}{rll} \text{minimize} & \pi z_2 - \pi z_1 & \\ \text{subject to} & (x_1 - a_i)^2 + (x_2 - b_i)^2 - z_1 \ \geq \ 0, & i = 1, \ldots, m, \\ & (x_1 - a_i)^2 + (x_2 - b_i)^2 - z_2 \ \leq \ 0, & i = 1, \ldots, m. \end{array}$$

5.(c) For each given point $(x_1, x_2)^{\mathsf{T}} \in I\!R^2$, the corresponding optimal values of z_1 and z_2 in the above problem are clearly

$\hat{z}_1(x_1, x_2) = \min_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\}$ and
$\hat{z}_1(x_1, x_2) = \max_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\}$.

Thus, the above problem can be written

$$\text{minimize} \quad \pi \max_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\} - \pi \min_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\},$$

which is a problem in just the two variables x_1 and x_2.
(However, the objective function is not differentiable, so this is not a correct answer to 5.(b).)
But $\min_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\} = \min_i \{x_1^2 - 2a_i x_1 + a_i^2 + x_2^2 - 2b_i x_2 + b_i^2\} =$
$x_1^2 + x_2^2 + \min_i \{-2a_i x_1 + a_i^2 - 2b_i x_2 + b_i^2\}$, and
$\max_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\} = \max_i \{x_1^2 - 2a_i x_1 + a_i^2 + x_2^2 - 2b_i x_2 + b_i^2\} =$
$x_1^2 + x_2^2 + \max_i \{-2a_i x_1 + a_i^2 - 2b_i x_2 + b_i^2\}$.

Therefore, $\max_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\} - \min_i \{(x_1 - a_i)^2 + (x_2 - b_i)^2\} =$
$\max_i \{-2a_i x_1 + a_i^2 - 2b_i x_2 + b_i^2\} - \min_i \{-2a_i x_1 + a_i^2 - 2b_i x_2 + b_i^2\}$,

so the above problem can be written

$$\text{minimize} \quad \pi \max_i \{-2a_i x_1 + a_i^2 - 2b_i x_2 + b_i^2\} - \pi \min_i \{-2a_i x_1 + a_i^2 - 2b_i x_2 + b_i^2\},$$

which may equivalently be written

$$\begin{array}{rll} \text{minimize} & \pi w_2 - \pi w_1 & \\ \text{subject to} & w_1 + 2a_i x_1 + 2b_i x_2 \ \leq \ a_i^2 + b_i^2, & i = 1, \ldots, m, \\ & w_2 + 2a_i x_1 + 2b_i x_2 \ \geq \ a_i^2 + b_i^2, & i = 1, \ldots, m, \end{array}$$

which is an LP problem in the variables x_1, x_2, w_1 and w_2.