Solutions to exam in SF1811 Optimization, 18 Jan 2014

1.(a) Let $\mathbf{x} = (\, x_{12},\, x_{13},\, x_{14},\, x_{23},\, x_{25},\, x_{34},\, x_{35},\, x_{45}\,)^{\mathsf{T}}$,
where the variable x_{ij} stands for the flow in the arc from node i to node j.
Let $\mathbf{c} = (\, c_{12},\, c_{13},\, c_{14},\, c_{23},\, c_{25},\, c_{34},\, c_{35},\, c_{45}\,)^{\mathsf{T}} = (\,1,\, k,\, k,\, 1,\, k,\, 1,\, k,\, 1)^{\mathsf{T}}$.
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

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 30 \\ 20 \\ 0 \\ -35 \\ -15 \end{pmatrix}.$$

Finally, the given directions of the arcs 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) If we let x_{12}, x_{23}, x_{34} and x_{45} be basic variables, the values of these basic variables can be calculated as follows:
$x_{12} = 30$, because of flow balance in node 1,
$x_{23} = 50$, because of flow balance in node 2,
$x_{34} = 50$, because of flow balance in node 3,
$x_{45} = 15$, because of flow balance in node 4.
We see that the flow balance condition in node 5 also becomes fulfilled (since the problem is balanced).

1.(c) The reduced costs for the nonbasic variables can be calculted by
$r_{ij} = c_{ij} - y_i + y_j$ for all nonbasic arcs,
where the scalars (simplex multipliers) y_i are calculated by
$y_i - y_j = c_{ij}$ for all basic arcs, and $y_5 = 0$.
We get:
$y_5 = 0$, (by definition)
$y_4 = c_{45} + y_5 = 1 + 0 = 1$,
$y_3 = c_{34} + y_4 = 1 + 1 = 2$,
$y_2 = c_{23} + y_3 = 1 + 2 = 3$,
$y_1 = c_{12} + y_2 = 1 + 3 = 4$,
and then
$r_{13} = c_{13} - y_1 + y_3 = k - 4 + 2 = k - 2$,
$r_{14} = c_{14} - y_1 + y_4 = k - 4 + 1 = k - 3$,
$r_{25} = c_{25} - y_2 + y_5 = k - 3 + 0 = k - 3$,
$r_{35} = c_{35} - y_3 + y_5 = k - 2 + 0 = k - 2$.
We se that if $k \geq 3$ then all $r_{ij} \geq 0$ and the given basic solution is optimal.
If $k < 3$ then $r_{14} < 0$ and then we could let
$x_{14} = t$, $x_{12} = 30-t$, $x_{23} = 50-t$, $x_{34} = 50-t$ and $x_{45} = 15$.
For $t \in (0, 30\,]$, this is a feasible solution with strictly decreasing cost when t increases. Thus, the basic solution from (b) is optimal if and only if $k \geq 3$.

2.(a) The considered LP problem is on the form

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

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

In the suggested solution, x_1 and x_2 are basic variables, i.e. $\beta = (1, 2)$ and $\nu = (3, 4)$.

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

The current values of the basic variables are $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where $\bar{\mathbf{b}}$ is obtained from

$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$,

and thus $x_1 = 3$ and $x_2 = 1$, which agrees with the suggested solution.

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

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

The reduced costs for the non-basic variables are given by

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

Since these reduced costs are non-negative, the suggested solution is optimal.
Thus, $\mathbf{x} = (3,\ 1,\ 0,\ 0)^\mathsf{T}$ is an optimal solution. The optimal value $= \mathbf{c}^\mathsf{T}\mathbf{x} = 6$.

2.(b) Since the primal problem is on the form

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

the corresponding dual problem D1 is on the form

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

which, written out in details, becomes

$$\begin{array}{ll} \text{maximize} & 2y_1 + 4y_2 \\ \text{subject to} & \begin{array}{rcl} y_1 + y_2 & \leq & 1, \\ -y_1 + y_2 & \leq & 3, \\ y_1 - y_2 & \leq & 1, \\ -y_1 - y_2 & \leq & 1. \end{array} \end{array}$$

A careful figure shows that the feasible region is a rectangle with corners $(1, 0)$, $(-1, 2)$, $(-2, 1)$ and $(0, -1)$. Level sets to the objective function $2y_1 + 4y_2$ are parallel lines orthogonal to the vector $(2, 4)$, with increasing values when moving "north-north-east".

The level set which corresponds to the maximal value of the objective function is given, from the figure, by the line $2y_1 + 4y_2 = 6$, which goes through the corner $(y_1, y_2) = (-1, 2)$. Thus, this is the optimal solution to the dual problem D1. The optimal value $= \mathbf{b}^\mathsf{T}\mathbf{y} = 6$.

2.(c) With surplus variables x_5 and x_6, an LP problem on standard form is obtained:

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

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

We start from the solution from (a), with $\beta = (1,2)$ and $\nu = (3,4,5,6)$,
which is a feasible basic solution also to this new problem.

The matrix $\mathbf{A}_\beta = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and the vectors $\bar{\mathbf{b}} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ are the same as in (a), while the matrix $\mathbf{A}_\nu = \begin{bmatrix} 1 & -1 & -1 & 0 \\ -1 & -1 & 0 & -1 \end{bmatrix}$.

The reduced costs for the non-basic varibles are now give by

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

Since $r_{\nu_3} = r_5 = -1$ is smallest, and < 0, we let x_5 become a new basic variable.

The vector $\bar{\mathbf{a}}_5$ is obtained from the system $\mathbf{A}_\beta \bar{\mathbf{a}}_5 = \mathbf{a}_5$, which becomes

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{a}}_1 = \begin{pmatrix} \bar{a}_{15} \\ \bar{a}_{25} \end{pmatrix} = \begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix}.$$

The value of the new basic variable x_5 is given by

$$t^{\max} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{i5}} \mid \bar{a}_{i5} > 0 \right\} = \frac{1}{0.5} = \frac{\bar{b}_2}{\bar{a}_{25}}.$$

Here, the minimizing index is $i = 2$, so $x_{\beta_2} = x_2$ should leave the basis.

Now $\beta = (1,5)$ and $\nu = (2,3,4,6)$.

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

The current values of the basic variables are $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where $\bar{\mathbf{b}}$ is obtained from

$$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}, \text{ i.e. } \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \text{ with the solution } \bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

The vector $\mathbf{y}$ with simplex multipliers is obtained from the system

$$\mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta, \text{ i.e. } \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ with the solution } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The reduced costs for the non-basic variables are given by

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

Since these reduced costs are non-negative, the current basic solution is optimal.
Thus, $\mathbf{x} = (4,\, 0,\, 0,\, 0,\, 2,\, 0)^\mathsf{T}$ is an optimal solution. The optimal value $= \mathbf{c}^\mathsf{T}\mathbf{x} = 4$.

2.(d) With slack variables x_5 and x_6, an LP problem on standard form is obtained:

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

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

As recommended, we now start with $\beta = (5, 6)$ and $\nu = (1, 2, 3, 4)$.

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

The current values of the basic variables are $\mathbf{x}_\beta = \bar{\mathbf{b}}$, where $\bar{\mathbf{b}}$ is obtained from

$\mathbf{A}_\beta \bar{\mathbf{b}} = \mathbf{b}$, i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$, with the solution $\bar{\mathbf{b}} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$.

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

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

The reduced costs for the non-basic variables are given by

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

Since these reduced costs are non-negative, the current basic solution is optimal.
Thus, $\mathbf{x} = (0,\ 0,\ 0,\ 0,\ 2,\ 4)^\mathsf{T}$ is an optimal solution. The optimal value $= \mathbf{c}^\mathsf{T}\mathbf{x} = 0$.

3.(a) 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 $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{c} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

The matrix $\mathbf{H} = \mathbf{I}$ is positive definite, so we have a convex QP problem.

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

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 $x_2 = 3 - v$.
Thus, the complete set of solutions to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is given by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} v = \bar{\mathbf{x}} + \mathbf{z}\, v,$$

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

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

and the unique global optimal solution to the original problem is

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{z}\,\hat{v} = \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

3.(b) The Lagrange optimality conditions for the considered convex

QP problem are given by the system $\begin{array}{rcl} \mathbf{Hx} - \mathbf{A}^\mathsf{T}\mathbf{u} & = & -\mathbf{c} \\ \mathbf{Ax} & = & \mathbf{b} \end{array}$

The equations $\mathbf{Hx} - \mathbf{A}^\mathsf{T}\mathbf{u} = -\mathbf{c}$ are in our case equivalent to $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{u} - \mathbf{c}$.
If this is combined with the remaining equations $\mathbf{Ax} = \mathbf{b}$, we get that
$\mathbf{AA}^\mathsf{T}\mathbf{u} = \mathbf{Ac} + \mathbf{b}$, which in our case becomes

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, with the unique solution $\hat{\mathbf{u}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

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

is then given by $\hat{\mathbf{x}} = \mathbf{A}^\mathsf{T}\hat{\mathbf{u}} - \mathbf{c} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

Since $\mathbf{H}$ is positive definite, the Lagrange conditions are both necessary and sufficient for a global optimum, and thus $\hat{\mathbf{x}}$ is the unique global optimal solution to the considered QP problem. As expected, the obtained results in (a) and (b) agree.

3.(c)

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

Then the considered problem becomes: minimize $f(\mathbf{x})$ subject to $g_i(\mathbf{x}) \leq 0$ for $i = 1, 2$.

The KKT conditions for this problem become

(KKT–1) $\dfrac{\partial f}{\partial x_j} + y_1 \dfrac{\partial g_1}{\partial x_j} + y_2 \dfrac{\partial g_2}{\partial x_j} = 0$ for $j = 1, 2, 3$, i.e.
$x_1 - 1 + y_1 = 0$,
$x_2 - 1 + y_1 - y_2 = 0$,
$x_3 - 1 - y_2 = 0$,

(KKT–2) $g_i(\mathbf{x}) \leq 0$ for $i = 1, 2$, i.e.
$x_1 + x_2 - 1 \leq 0$,
$3 - x_2 - x_3 \leq 0$,

(KKT–3) $y_1 \geq 0$ och $y_2 \geq 0$.

(KKT–4) $y_i g_i(\mathbf{x}) = 0$ for $i = 1, 2$, i.e.
$y_1(x_1 + x_2 - 1) = 0$,
$y_2(3 - x_2 - x_3) = 0$.

Let $\hat{\mathbf{x}} = (0, 1, 2)^\mathsf{T}$, as in (a) and (b). Then $g_1(\hat{\mathbf{x}}) = 0$ and $g_2(\hat{\mathbf{x}}) = 0$ so that (KKT–2) and (KKT–4) are satisfied by $\hat{\mathbf{x}}$, for all y_1 and y_2.

Further, the conditions (KKT–1) are satisfied by $\hat{\mathbf{x}} = (0, 1, 2)^\mathsf{T}$ and $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2)^\mathsf{T}$ if and only if $\hat{\mathbf{y}} = (1, 1)^\mathsf{T}$. But $\hat{\mathbf{y}} = (1, 1)^\mathsf{T}$ satisfies also (KKT–3).

Thus, $\hat{\mathbf{x}} = (0, 1, 2)^\mathsf{T}$ is a KKT point. Since f, g_1 and g_2 are convex functions, every KKT point is a global optimal solution, and thus $\hat{\mathbf{x}}$ is a global optimal solution to the considered inequality-constrained QP problem.

3.(d)

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

The KKT conditions for this problem become

(KKT–1) $\dfrac{\partial f}{\partial x_j} + y_1 \dfrac{\partial g_1}{\partial x_j} + y_2 \dfrac{\partial g_2}{\partial x_j} = 0$ for $j = 1, 2, 3$, i.e.

$$x_1 - 1 - y_1 = 0,$$
$$x_2 - 1 - y_1 + y_2 = 0,$$
$$x_3 - 1 + y_2 = 0,$$

(KKT–2) $g_i(\mathbf{x}) \le 0$ for $i = 1, 2$, i.e.

$$1 - x_1 - x_2 \le 0,$$
$$x_2 + x_3 - 3 \le 0,$$

(KKT–3) $y_1 \ge 0$ och $y_2 \ge 0$.

(KKT–4) $y_i g_i(\mathbf{x}) = 0$ for $i = 1, 2$, i.e.

$$y_1(1 - x_1 - x_2) = 0,$$
$$y_2(x_2 + x_3 - 3) = 0.$$

The result from (c) indicate that the objective function would be decreased (compared to the solution in (c)) if $x_1 + x_2 > 1$ and $x_2 + x_3 < 3$, i.e. if the constraints in (d) are not satisfied with equality. So let us try with $y_1 = y_2 = 0$ in the KKT conditions above.

Then (KKT–3) and (KKT–4) are satisfied for all $\mathbf{x}$.

Further, the conditions (KKT–1) are satisfied if and only if $\mathbf{x} = (1, 1, 1)^\mathsf{T}$.
But this $\mathbf{x}$ satisfies also (KKT–2)!

Thus, $\mathbf{x} = (1, 1, 1)^\mathsf{T}$ is a KKT point. Again, since f, g_1 and g_2 are convex functions, every KKT point is a global optimal solution, and thus $\mathbf{x} = (1, 1, 1)^\mathsf{T}$ is a global optimal solution.

Interpretation (and a shorter way to solve the problem): If the constraints are completely neglected, so that the problem becomes simply to minimize $\frac{1}{2}(x_1^2 + x_2^2 + x_3^2) - x_1 - x_2 - x_3$ without any constraints, then the unique optimal solution would clearly be $\mathbf{x} = (1, 1, 1)^\mathsf{T}$. But since this solution happens to satisfy the constraints in (d), it must be the unique optimal solution also to the problem in (d). (Note that $\mathbf{x} = (1, 1, 1)^\mathsf{T}$ is not feasible, and thus not optimal, to the problems in (a)–(c).)

4.

Change notation on the constant from c to x.

Then we should minimize $f(x) = \frac{1}{2}\,\mathbf{h}(x)^\mathsf{T}\mathbf{h}(x) = \frac{1}{2}(h_1(x)^2 + h_2(x)^2)$,

where $\mathbf{h}(x) = \begin{pmatrix} h_1(x) \\ h_2(x) \end{pmatrix}$, with

$$h_1(x) = \frac{1}{1 + x\,t_1} - w_1 = \frac{1}{1+x} - 0.46,$$

$$h_2(x) = \frac{1}{1 + x\,t_2} - w_2 = \frac{1}{1+3x} - 0.22.$$

This is a nonlinear least-squares problem with $n = 1$ (one single variable x) and $m = 2$ (two terms in the quadratic sum).

Differentiation gives that

$$\nabla\mathbf{h}(x) = \begin{bmatrix} h_1'(x) \\ h_2'(x) \end{bmatrix}, \text{ where } h_1'(x) = \frac{-1}{(1+x)^2} \text{ and } h_2'(x) = \frac{-3}{(1+3x)^2}.$$

We should start in $x^{(1)} = 1$. Then

$$\mathbf{h}(x^{(1)}) = \begin{pmatrix} 0.04 \\ 0.03 \end{pmatrix} \text{ and } \nabla\mathbf{h}(x^{(1)}) = \begin{bmatrix} -1/4 \\ -3/16 \end{bmatrix} = \begin{bmatrix} -4/16 \\ -3/16 \end{bmatrix}.$$

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

In our case, $\nabla\mathbf{h}(x^{(1)})^\mathsf{T}\nabla\mathbf{h}(x^{(1)}) = (-4/16)^2 + (-3/16)^2 = 25/256$

and $-\nabla\mathbf{h}(x^{(1)})^\mathsf{T}\mathbf{h}(x^{(1)}) = (4/16)(4/100) + (3/16)(3/100) = 25/1600$,

so we get the equation $(25/256)\,d = 25/1600$, with the solution $d^{(1)} = 256/1600 = 0.16$.

We try with $t_1 = 1$, so that $x^{(2)} = x^{(1)} + t_1 d^{(1)} = 1 + 0.16 = 1.16$. Then

$$h_1(x^{(2)}) = \frac{1}{2.16} - 0.46 = \frac{1 - 0.46 \cdot 2.16}{2.16} = \frac{0.0064}{2.16} < 0.04 = h_1(x^{(1)}) \text{ and}$$

$$h_2(x^{(2)}) = \frac{1}{4.48} - 0.22 = \frac{1 - 0.22 \cdot 4.48}{4.48} = \frac{0.0144}{4.48} < 0.03 = h_2(x^{(1)}).$$

Since $|h_1(x^{(2)})| < |h_1(x^{(1)})|$ and $|h_2(x^{(2)})| < |h_2(x^{(1)})|$ it follows that $f(x^{(2)}) < f(x^{(1)})$, which means that we should accept $t_1 = 1$.

Now we have made an iteration with Gauss-Newtons method and obtained the new suggested value $c = 1.16$, which is better than the starting value $c = 1$ since the quadratic sum $\frac{1}{2}\sum_{i=1}^{m}(\frac{1}{1 + c\,t_i} - w_i)^2$ has decreased.

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

$L(\mathbf{x}, y) = (\mathbf{x}-\mathbf{q})^\mathsf{T}(\mathbf{x}-\mathbf{q}) + y\cdot(\mathbf{x}^\mathsf{T}\mathbf{D}\mathbf{x} - 1)$, with $\mathbf{x} \in I\!R^n$ and $y \in I\!R$.

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

Since $L(\mathbf{x}, y) = \mathbf{x}^\mathsf{T}(\mathbf{I} + y\,\mathbf{D})\mathbf{x} - 2\mathbf{q}^\mathsf{T}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{q} - y$, the optimal solution to PR_y is given by

$\tilde{\mathbf{x}}(y) = (\mathbf{I} + y\,\mathbf{D})^{-1}\mathbf{q}$, i.e. $\tilde{x}_j(y) = \dfrac{q_j}{1 + y\,d_j}$, for $j = 1, \ldots, n$.

Then the dual objective function becomes

$$\varphi(y) = L(\tilde{\mathbf{x}}(y), y) = -\mathbf{q}^\mathsf{T}(\mathbf{I} + y\,\mathbf{D})^{-1}\mathbf{q} + \mathbf{q}^\mathsf{T}\mathbf{q} - y = \mathbf{q}^\mathsf{T}\mathbf{q} - y - \sum_{j=1}^{n} \frac{q_j^2}{1 + y\,d_j}\,.$$

The dual problem consists of maximizing $\varphi(y)$ with respect to $y \geq 0$.

5.(b) Some calculus give that

$$\varphi'(y) = -1 + \sum_{j=1}^{n} \frac{q_j^2 d_j}{(1 + y\,d_j)^2} \quad \text{and} \quad \varphi''(y) = -\sum_{j=1}^{n} \frac{2q_j^2 d_j^2}{(1 + y\,d_j)^3}\,.$$

In particular, $\varphi'(0) = -1 + \sum_{j=1}^{n} q_j^2 d_j = \mathbf{q}^\mathsf{T}\mathbf{D}\,\mathbf{q} - 1 > 0$.

Further, $\varphi''(y) < 0$ for all $y \geq 0$, which implies that $\varphi'(y)$ is continuous and strictly decreasing for $y \geq 0$, and also that $\varphi(y)$ is strictly concave for $y \geq 0$.

Finally, $\varphi'(y_1) < -1 + \sum_{j=1}^{n} \dfrac{q_j^2 d_j}{(y_1 d_j)^2} = -1 + \dfrac{1}{y_1^2}\sum_{j=1}^{n} \dfrac{q_j^2}{d_j} = 0$ if $y_1^2 = \sum_{j=1}^{n} \dfrac{q_j^2}{d_j} > 0$.

5.(c) The results in (b) imply that there is a unique $\hat{y} > 0$ such that $\varphi'(\hat{y}) = 0$. In addition, since $\varphi(y)$ is strictly concave for $y \geq 0$, this unique $\hat{y} > 0$ which satisfies $\varphi'(\hat{y}) = 0$ is the unique optimal solution to the dual problem.

Now let $\hat{\mathbf{x}} = \tilde{\mathbf{x}}(\hat{y})$. Then $\hat{x}_j = \tilde{x}_j(\hat{y}) = \dfrac{q_j}{1 + \hat{y}\,d_j}$, so that $\varphi'(\hat{y}) = 0$ implies that

$$0 = \sum_{j=1}^{n} \frac{q_j^2 d_j}{(1 + \hat{y}\,d_j)^2} - 1 = \sum_{j=1}^{n} d_j \hat{x}_j^2 - 1 = \hat{\mathbf{x}}^\mathsf{T}\mathbf{D}\,\hat{\mathbf{x}} - 1.$$

Then $\hat{\mathbf{x}}$, together with $\hat{y}$, satisfies the global optimality conditions (GOC):

(i) $L(\hat{\mathbf{x}}, \hat{y}) \leq L(\mathbf{x}, \hat{y})$ for all $\mathbf{x} \in I\!R^n$. (Since $\hat{\mathbf{x}} = \tilde{\mathbf{x}}(\hat{y})$.)

(ii) $\hat{\mathbf{x}}^\mathsf{T}\mathbf{D}\,\hat{\mathbf{x}} - 1 \leq 0$. (Since $\hat{\mathbf{x}}^\mathsf{T}\mathbf{D}\,\hat{\mathbf{x}} = 1$.)

(iii) $\hat{y} \geq 0$. (Since $\hat{y} > 0$.)

(iv) $\hat{y}\cdot(\hat{\mathbf{x}}^\mathsf{T}\mathbf{D}\,\hat{\mathbf{x}} - 1) = 0$. (Since $\hat{\mathbf{x}}^\mathsf{T}\mathbf{D}\,\hat{\mathbf{x}} = 1$.)

This implies that $\hat{\mathbf{x}}$ is a global optimal solution to P.